Molecular Dynamics Simulation of Parvalbumin in Aqueous Solution

Peter Ahlström,* Olle Teleman, Bo Jönsson, and Sture Forsén

Contribution from the Department of Physical Chemistry 2, Chemical Centre, University of Lund, S-221 00 Lund, Sweden. *Received May 27, 1986*

Abstract: Molecular dynamics simulations of the Ca^{2+} binding protein parvalbumin are reported. The calcium-saturated protein is simulated in an aqueous environment but also in vacuo. The Ca^{2+} free protein is studied as well, albeit only in vacuo. An X-ray diffraction structure is the starting point for all three simulations. Structural and dynamic properties are investigated and compared between the different simulations, each which lasted ≈100 ps. The two in vacuo simulations show considerable structural changes relative to the initial coordinates. The entire protein contracts, and its helix structures change. When simulated in aqueous solution, the protein still changes its structure relative to the crystal form, but less so than in vacuo, and helix structures largely remain intact. Dynamic properties of the protein are considerably altered on inclusion of water. Reorientational correlation times for vectors, both in the backbone and in side chains, change by orders of magnitude. This is the case for surface residues as well as internal. The calcium binding sites are accessible to the solvent, and ligand exchanges are seen during the simulation. Removal of calcium ions causes global structural and dynamic changes. Internal protein hydrogen bonds are investigated. The number of hydrogen bonds is approximately the same in all three simulations, whereas their lifetime seems to be slightly longer in the in vacuo simulations. In general they are short lived, with an average lifetime of a few picoseconds.

During the last decade dynamic properties of biological macromolecules, in particular proteins and nucleic acids, have received considerable attention.[1] This development has been stimulated by the recognition that the image of proteins and nucleic acids as rigid structures is highly incomplete and fails to capture the dynamic nature of biological processes. The binding of a substrate molecule to an enzyme, the chemical transformation of the substrate, and its subsequent release must be associated with considerable conformational adjustments and changes in the mobility of many atoms. A much quoted example concerns the binding of an oxygen molecule to myoglobin or hemoglobin—the equilibrium structures of the deoxy form of these molecules leave no room for the diatomic ligand to reach the buried heme binding sites.[2-4] Although our present understanding of molecular mechanisms of biological processes is far from complete, it appears that internal motions of proteins, relevant to their function, involve many degrees of freedom and cover a broad time scale ranging from picoseconds to seconds.

Experimental information regarding internal motions in proteins has been obtained through several methods. An early method, pioneered by Linderström-Lang,[5] is the exchange of hydrogen atoms between NH and OH groups of the proteins and solvent water. With the use of ^1H NMR methods, detailed information on the role of exchange of individual groups has been obtained in some cases, although the molecular details of the steps leading to exchange still remain obscure.[6-10] Particularly relevant are single-crystal neutron diffraction studies, which show that most internal amide protons in a crystalline protein can be exchanged with overall rates comparable to those for the protein in solution.[11-15] Studies of the dynamic quenching of fluorescent amino acids by oxygen and other molecules[16-18] have indicated the interior of proteins to be accessible to small molecules.

Time-resolved fluorescent anisotropy studies have been used to determine the motion of tyrosine and tryptophan side chains. Rotational correlation times in the order of a few hundred picoseconds or less have been measured.[19,20] Perhaps the most detailed experimental information on internal motions in proteins has been provided by NMR spectroscopy. Recent developments in two-dimensional Fourier transform NMR have considerably simplified the assignment of observed resonances to atoms in individual amino acids in proteins. Measurements of relaxation rates (longitudinal and transverse) and nuclear Overhauser enhancements (NOE) probe internal motions in the time range from microseconds to fractions of nanoseconds. In addition to the dynamic information obtainable from the techniques mentioned above, information on the mean square fluctuations of atoms from their equilibrium positions in crystalline proteins may be found from Debye–Waller factors.[21] In the case of heme proteins, supplementary information on lattice disorder effects has been obtained from Mössbauer data.[4,21]

A fundamental problem, common to all experimental techniques mentioned, concerns the detailed molecular nature of the dynamic processes observed. This particular problem is not inherent in some of the theoretical approaches to macromolecular internal mobility. The dynamics of a protein molecule can in principle be modeled from statistical mechanical simulation techniques. Provided that the potential energy of the macromolecule is known as a function of the positions of all constituent atoms, the trajectories of all atoms are obtained by numerical integration of Newton's equations of motion. Starting with the pioneering molecular dynamics (MD) simulation by McCammon, Gelin, and Karplus of the bovine pancreatic trypsin inhibitor (BPTI) with 58 amino acids,[22,23] this theoretical approach has attracted considerable attention. For a review of the application of molecular dynamics simulations and related methods to proteins, see, for example, ref 24 and 25.

(1) Frauenfelder, H. *Ciba Found. Symp.* **1983**, *93*, 329.
(2) Perutz, M. F.; Mathews, F. S. *J. Mol. Biol.* **1966**, *21*, 199.
(3) Takano, T. *J. Mol. Biol.* **1977**, *110*, 569.
(4) Debrunner, P. G.; Frauenfelder, H. *Annu. Rev. Phys. Chem.* **1982**, *33*, 283.
(5) Linderström-Lang, K. *Spec. Publ.—Chem. Soc.* **1955**, *2*, 1.
(6) Hvidt, A.; Nielsen, S. O. *Adv. Protein Chem.* **1966**, *21*, 287.
(7) Wagner, G. *Q. Rev. Biophys.* **1983**, *16*, 1.
(8) Allewell, N. M. *J. Biochem. Biophys. Methods* **1983**, *7*, 345.
(9) Englander, S. W.; Kaltenbach, N. *Q. Rev. Biophys.* **1984**, *16*, 521.
(10) Woodward, C. K.; Hilton, B. D. *Biophys. J.* **1980**, *32*, 561.
(11) Praissman, M.; Rupley, J. A. *Biochemistry* **1968**, *7*, 2446.
(12) Tüchsen, E.; Hvidt, A.; Ottesen, M. *Biochimie* **1980**, *62*, 563.
(13) Hanson, J.; Schoenborn, B. *J. Mol. Biol.* **1981**, *153*, 117.
(14) Kossiakoff, A. A. *Nature (London)* **1982**, *296*, 713.
(15) Wlodauer, A.; Sjölin, L. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *80*, 3628.
(16) Lakowicz, J. R.; Weber, G. *Biochemistry* **1973**, *12*, 4171.
(17) Eftink, M. R.; Ghiron, C. A. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 3290.
(18) Eftink, M. R.; Ghiron, C. A. *Biochemistry* **1977**, *16*, 5546.
(19) Munro, I.; Pecht, I.; Stryer, L. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 56.
(20) Hochstrasser, R. M.; Negus, D. K. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 4399.
(21) Petsko, G. A.; Ringe, D. *Annu. Rev. Biophys. Bioeng.* **1984**, *13*, 331.
(22) McCammon, J. A.; Gelin, B. R.; Karplus, M. *Nature (London)* **1977**, *267*, 585.
(23) Karplus, M.; McCammon, J. A. *Nature (London)* 1979, *277*, 578.

0002-7863/87/1509-1541$01.50/0

Most MD simulations of biological macromolecules have been done without explicit consideration of solvent molecules. Simulations of BPTI in a van der Waals solvent, the molecules of which having the density and molecular size corresponding to water molecules, have been made by van Gunsteren and Karplus[26] and by Swaminathan et al.[27] van Gunsteren et al.[28] performed a short MD simulation of a unit cell of the BPTI crystal involving 4 protein molecules and 560 water molecules. The first short (20-ps) simulation of BPTI surrounded by dipolar water molecules was made by van Gunsteren and Berendsen[29] in 1984. The magnitudes of the fluctuations calculated agree reasonably well with those calculated earlier for BPTI in vacuo, whereas the average structure was closer to the X-ray results. Recently, a study by Krüger et al. of avian pancreatic polypeptide hormone in aqueous solution as well as in crystal[30] was published. All these studies have been taken to indicate that, while the motion of side chains at the protein surface will not be calculated correctly, solvent effects on fluctuations in the interior of proteins will generally be small.

In the present work we have attempted to investigate more closely the effects of solvent water on the dynamic properties of a Ca^{2+} binding protein. Parallel MD runs have been made in vacuo and in water. Time correlation functions pertaining to different internal motions have been compared as have been radii of gyration, R factors, rotational diffusion, and other properties.

The protein chosen for the present study, parvalbumin (M_r = 11 471, see Figure 1), belongs to a class of Ca^{2+} binding proteins that plays an important role in living systems.[31-33] This class includes calmodulin, a ubiquitous protein vital to the Ca^{2+}-dependent regulation of a wide variety of cellular events, and troponin C, a protein present in all skeletal muscles and responsible for the Ca^{2+}-dependent triggering of muscle contraction. These proteins are structurally homologous, and the Ca^{2+} binding sites all feature the same helix–loop–helix structural arrangement—usually termed the "EF hand".[32,34] Parvalbumin has two EF hand type sites and is capable of binding two Ca^{2+} ions with high affinity ($K_A = 10^9\ M^{-1}$).[33] This protein therefore also offers an opportunity not earlier explored in MD simulations, namely, to study the effect of removal of intrinsic metal ions on the dynamic properties of a protein. To investigate this a third MD simulation of parvalbumin with the two Ca^{2+} ions removed has been performed in vacuo.

Methods

Interaction Potentials. The inter- and intramolecular interactions in a protein solution are immensely complicated and to a large extent inaccessible to quantum chemical calculations. Ab initio quantum chemical calculations yield intermolecular interactions between small molecules and ions as well as the internal force field of small molecules. For example, the interaction between two water molecules has been calculated by Matsuoka et al.,[35] and the resulting pair potential has been used in subsequent simulations of the liquid.[36] Quantum chemical calculations of intermolecular potentials between water and amino acids have been reported by Clementi and co-workers.[37-39] These potentials can, after minor manipulations, be combined so as to give the interaction between a water molecule and a peptide or a protein.[40] However, the accuracy of such a potential is difficult to assess. For example, the internal flexibility in the amino acid residues is not taken into account properly. This is the present limit of quantum chemical techniques, and the interaction within a protein or alternatively between different amino acids is still beyond reach.

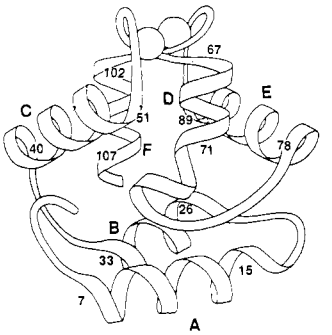

Figure 1. Schematic drawing of the X-ray structure of the backbone of carp parvalbumin. The helices are named A–F according to standard notation, and numbers refer to amino acid numbers in the primary structure. On top are the two calcium binding sites (after an original drawing by J. Richardson).

In order to simulate protein solutions one has to resort to empirical or semiempirical potentials. A number of different force fields have been proposed[41-45] and used in protein simulations. The accuracy and validity of these empirical potentials are in general unknown. In most cases it is very difficult to find appropriate experimental data with which to test a potential. The confidence in the empirical potentials probably stems more from the analogy with simple liquids than from rigorous tests. An interesting question to ask is how accurate potentials we do need. There is no straightforward answer, since it depends on what properties we want to study. From our point of view the intermolecular potentials used hitherto seem to be rather crude, and the protein simulations must be considered to be under development. This may sometimes be forgotten, and results may not be interpreted with adequate caution.

So far, most protein simulations have been carried out in vacuo, and the interest has been focused on small globular proteins with a minimum of charged residues. In reality, however, proteins do occur in solution and have a fair amount of charged residues, which may give rise to a considerable net charge. We believe that electrostatic interactions play a vital role in biological processes, and the theoretical and technical problems associated therewith have to be dealt with. In many in vacuo simulations reported, the electrostatic interactions are attenuated by assuming a "distance dependent dielectric permittivity", which is supposed to mimick the solvent.[45] The theoretical justification for such a procedure is not immediate, and it has been subject to criticism.[41] In a true molecular model, screening of electrostatic interactions should of course be due to the solvent. This places severe requirements on the dielectric behavior of the water model used, and it may preclude pair-wise additivity. One argument for the relevance of protein simulations in vacuo seems to be that, although the amino acid residues at the surface will behave differently with

(24) McCammon, J. A.; Karplus, M. *Acc. Chem. Res.* **1983**, *16*, 187.
(25) McCammon, J. A. *Rep. Prog. Phys.* **1984**, *47*, 1.
(26) van Gunsteren, W. F.; Karplus, M. *Biochemistry* **1982**, *21*, 2259.
(27) Swaminathan, S.; Ichiye, T.; van Gunsteren, W.; Karplus, M. *Biochemistry* **1982**, *21*, 5230.
(28) van Gunsteren, W. F.; Berendsen, H. J. C.; Hermans, J.; Hol, W. G. J.; Postma, J. P. M. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 4315.
(29) van Gunsteren, W. F.; Berendsen, H. J. C. *J. Mol. Biol.* **1984**, *176*, 559.
(30) Krüger, P.; Strassburger, W.; Wollmer, A.; van Gunsteren, W. F. *Eur. Biophys. J.* **1985**, *13*, 77.
(31) Wnuk, W.; Cox, J. A.; Stein, E. A. In *Calcium and Cell Function*; Cheung, W. Y., Ed.; Academic: New York, 1982; Vol. II, p 243.
(32) Demaille, J. G. In *Calcium and Cell Function*; Cheung, W. Y., Ed.; Academic: New York, 1982; Vol. II, p 111.
(33) Seamon, K. B.; Kretsinger, R. H. In *Calcium in Biology*; Spiro, T. G., Ed.; Wiley-Interscience: New York, 1983; p 1.
(34) Kretsinger, R. H. *C.R.C. Crit. Rev. Biochem.* **1980**, *8*, 119.
(35) Matsuoka, O.; Clementi, E.; Yoshimine, M. *J. Chem. Phys.* **1976**, *64*, 1351.
(36) Lie, G. C.; Clementi, E.; Yoshimine, M. *J. Chem. Phys.* **1976**, *64*, 2314.
(37) Clementi, E.; Cavallone, F.; Scordamaglia, R. *J. Am. Chem. Soc.* **1977**, *99*, 5531.
(38) Bolis, G.; Clementi, E. *J. Am. Chem. Soc.* **1977**, *99*, 5550.
(39) Clementi, E.; Corongiu, G.; Ranghino, G. *J. Chem. Phys.* **1981**, *74*, 578.
(40) Clementi, E.; Corongiu, G.; Jönsson, B.; Romano, S. *FEBS Lett.* **1979**, *100*, 313.
(41) Lifson, S. *NATO Adv. Study Inst./FEBS Adv. Course No. 78*, 1981.
(42) Levitt, M.; Lifson, S. *J. Mol. Biol.* **1969**, *46*, 269.
(43) Hagler, A. T.; Lapiccirella, A. *Pept.: Chem., Struct. Biol., Proc. Am. Pept. Symp., 4th* **1975**, 279.
(44) Hermans, J.; Berendsen, H. J. C.; van Gunsteren, W. F.; Postma, J. P. M. *Biopolymers* **1984**, *23*, 1513.
(45) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comp. Chem.* **1983**, *4*, 187.

or without solvent, the interior should not be significantly affected. The validity of this assumption has to our knowledge not yet been verified.

In an attempt to approach these problems we have constructed an inter- and intramolecular potential based on site–site interactions, where a site can be either an atom or a group of atoms (pseudoatom). The only pseudoatoms are CH, CH_2, and CH_3 groups, and remaining hydrogen atoms have been treated explicitly. This leads to considerable savings in computing time for in vacuo simulations. However, when simulating a protein in aqueous solution, the aliphatic hydrogen atoms only represent a small fraction of the total number of atoms. Thus they can be included at only a marginal increase in computational efforts. In fact, the aliphatic hydrogen atoms may play a significant role both for the protein structure and dynamics. For example, there is a nonnegligible dipole moment, ≈0.4 D, associated with a C–H bond. In the present simulations we have refrained from using explicit hydrogen atoms in order to make more direct comparisons with previous simulation works.

A pair of sites i and j belonging to different molecules or separated by more than two bonds interact via a Lennard–Jones (LJ) plus a Coulomb potential (eq 1). ϵ_{ij} and σ_{ij} are the usual

$$U_{ij}(r) = 4\epsilon_{ij}[(\sigma_{ij}/r)^{12} - (\sigma_{ij}/r)^{6}] + q_iq_j/4\pi\epsilon_o\epsilon_r r_{ij} \quad (1)$$

LJ parameters, and q_i is the partial charge on site i. Equation 1 is the standard expression for the intermolecular potential in protein simulations. The only, but important, difference is that ϵ_r is no longer distance dependent but is set equal to 1. The parameters have been taken from the literature,[44,46-48] and the LJ parameters were calculated from the Kirkwood–Slater formula[47] and from the assumption that the LJ potential minimum corresponds to the sum of the van der Waals radii of the two atoms. The LJ parameters chosen may not be the most accurate but should hopefully give a reasonable description of short-ranged interactions in the system. For example, the H_2O–amino acid potentials of Clementi and co-workers may well be a better choice than the now available empirical ones. For the present simulations, we have chosen not to use the quantum chemical potentials of Clementi et al., as their implementation is somewhat involved.

For the water–water interaction a number of different models exist, which all reproduce simple thermodynamic properties such as pair correlation functions, pressure, and internal energy with reasonable accuracy. In the present work we have used a modification of the empirical simple point charge (SPC) model of Berendsen et al.[48] The main reason for this choice was simplicity, as the SPC model conforms to the interaction expression, eq 1. Unfortunately, there has not, to our knowledge, been any determination of the dielectric permittivity for SPC water. One such simulation for the nonempirical potential of Matsuoka et al.[35] was reported recently, and its dielectric permittivity was found to be about 35 at room temperature,[52] roughly half the experimental value for water. One likely reason for the too low value is the assumption of pairwise additivity, neglecting nuclear and electronic polarization. One may argue that SPC water should behave better, since it is an effective pair potential including many-body polarization terms in an average way. However, this may not be the case, since the nonempirical potential of Matsuoka et al. has been fitted to a site–site model with a dipole moment of about 2.2 D, considerably larger than the experimental gas-phase value and close to the dipole moment of a SPC water molecule. Thus we expect SPC water to give a too low screening of ionic interactions.

Table I. Simulation Parameters and Averages

	APO	VAC	AQ
number of atoms	983	985	7970
box size, Å	∞	∞	43 × 50 × 43
time steps $\Delta t/\Delta T$, fs	0.2/1.0	0.2/1.0	0.2/1.2
equilibration time, ps	29.0	29.0	36.0
length of trajectory, ps	80.0	80.0	70.0
velocity scaling interval, fs	200	200	99.6
average scaling factor	0.97	0.96	0.98
neighbor list update interval, fs	12	12	4.8
CPU time/ps, s[a,b]	479	479	2523
CPU time/pair interaction, μs[a]	5.7	5.7	2.9
average temperature, K	305	306	303
average pressure, MPa			143

[a] The difference in timing between the in vacuo and solution simulations partly reflects improvements in the MD program and in the compiler. [b] All simulations were performed on a Cray-1A, and timing refers to total CPU time.

The internal bond and bond angle vibrations were treated explicitly, and assuming harmonicity the potential functions are given by eq 2, where x_i and $x_{i,eq}$ are the actual bond length and

$$U_{intra} = \sum_{bonds} K_b(x_i - x_{i,eq})^2/2 + \sum_{angles} K_a(\theta_i - \theta_{i,eq})^2/2 \quad (2)$$

equilibrium distance, respectively, with similar notation for angles. The force constants K_a and K_b and the equilibrium values were taken from the literature.[46,48-51] The interactions due to internal rotation were handled by means of periodic dihedral (torsional) potentials in the usual way,[41] eq 3, where Ψ_k is a dihedral angle

$$U_{int,rot} = \sum_{k}\sum_{n=1}^{3} C_{k,n}(1 - \cos[n\Psi_k]) \quad (3)$$

and $C_{k,n}$ an interaction parameter; k runs over all dihedral angles at each bond. The parameters in eq 1–3 are available as supplementary material. Improper dihedrals (out-of-plane torsions) were not included explicitly but treated by means of bond angle potentials. No explicit hydrogen bond potential was used; instead, it is assumed to be adequately described by Coulomb and LJ terms.

Simulation Technique. The molecular dynamics program used has been described in detail elsewhere,[53] and here only a brief account of the technique will be given. The program is based on independent atoms or pseudoatoms and hence includes covalent interactions in terms of harmonic and periodic potentials. This usually requires a rather small time step when integrating Newton's equations of motion, due to rapid bond and bond angle vibrations. In order to circumvent this problem we have devised a double time step algorithm,[53] where rapidly varying degrees of freedom are integrated with a small time step, while for slowly varying ones a larger time step can be used. In practice, bond and bond angle vibrations are integrated with Δt = 0.2 fs (fs = 10^{-15} s) while $\Delta T \approx$ 1–2 fs is used for the remaining degrees of freedom. The two-step algorithm gives radial distribution functions identical with those of a one-step algorithm with a time step of 0.2 fs. In the present simulations less than 10% of the total computation time was spent calculating the rapid motions; this is thus an efficient way to handle flexible molecules. It turned out that the Gear algorithm[54] is particularly suited for multiple time steps, and we used a fourth-order version of it.

Our algorithm differs from the usual approach, where bond lengths are kept fixed at their equilibrium value by using constraints. There are several advantages with the double time step algorithm: (i) the computer program is more easily "vectorized", (ii) the algorithm is stable even with a very repulsive start configuration, and (iii) the polarization of bond lengths and angles

(46) van Gunsteren, W. F.; Karplus, M. *Macromolecules* **1982**, *15*, 1528.
(47) Margenau, M.; Kestner, N. R. *Theory of Intermolecular Forces*; Pergamon: New York, 1969.
(48) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, 1981.
(49) Dolphin, D.; Wick, A. E. *Tabulation of Infrared Spectral Data*; Wiley-Interscience: New York, 1977.
(50) Herzberg, G. *Molecular Spectra and Molecular Structure: Infrared and Raman Spectra of Polyatomic Molecules*; Van Nostrand: Princeton, 1945.
(51) The parameters for the Cys residue were inferred from experimental data about methanethiol and ethanethiol: Schegel, H. B.; Wolfe, S.; Bernardi, F. *J. Chem. Phys.* **1977**, *67*, 4181. Durig, J. R.; Bucy, W. E.; Wurrey, C. J.; Carreira, L. A. *J. Phys. Chem.* **1975**, *79*, 988.
(52) Neumann, M. *J. Chem. Phys.* **1985**, *82*, 5663.
(53) Teleman, O.; Jönsson, B. *J. Comput. Chem.* **1986**, *7*, 58.
(54) Gear, C. W. *Numerical Initial Value Problems in Ordinary Differential Equations*; Prentice-Hall: Englewood Cliffs, NJ, 1971; p 148.

is included, although only classically. The latter point may be important for the dielectric properties of the solution. A more detailed discussion of the consequences of including intramolecular vibrations in water is found in ref 55.

Two in vacuo simulations were performed, one with Ca^{2+} ions at the two binding sites and one devoid of calcium. Below, these simulations will be referred to as the VAC and APO simulations, respectively. A third simulation, AQ, was performed in aqueous solution with 2327 water molecules and 3 sodium ions so as to make the whole sample electroneutral. This corresponds to a protein concentration of 18 mM. Technical details for the three simulations are summarized in Table I. All three simulations were started from the X-ray coordinates (denoted F6-A) of Kretsinger and Nockolds.[56] In the AQ simulation water molecules were placed on the bases of a primitive cubic lattice outside the protein. Three randomly chosen water molecules were exchanged for sodium.

Periodic boundary conditions were used in the AQ simulation, while the two others were performed in a cluster. All three simulations used a neighbor list technique and a spherical truncation of all interactions at $r_{cut} = 9$ Å, which leads to substantial computational savings in the AQ simulation. The neighbor list does not render itself entirely to vectorization but entails gather and scatter operations, which are quite time consuming. In consequence, a noncovalent potential (eq 1) of a more complex functional form only affects the total CPU time marginally.

Results and Discussion

Overall Rotation and Diffusion. Reorientational autocorrelation functions for vectors connecting two atoms have been calculated according to

$$C_j(t) = \langle P_j(\cos \theta(\vartheta,t))\rangle_\vartheta \tag{4}$$

where $\theta(\vartheta,t)$ is the angle between the vector at a time ϑ and the same vector at a time t later, P_j is the Legendre polynomial of order j, and the brackets indicate a time average. The jth correlation time, τ_j, is defined through

$$\tau_j = \int_0^\infty C_j(t)\, dt \tag{5}$$

If $C_j(t)$ decays as a single exponential, i.e.

$$C_j(t) = A \exp(-t/T) \tag{6}$$

then the correlation time τ_j equals the time constant T. Sometimes it is more appropriate to approximate the correlation function with a sum of exponentials as in eq 7, where the different time constants

$$C_j(t) = A \exp(-t/\tau_{j1}) + B \exp(-t/\tau_{j2}) + C \exp(-t/\tau_{j3}) + \ldots \tag{7}$$

may represent different processes, i.e., protein tumbling, side-chain motions, etc. There is often a rapid initial decay in a time correlation function, which is due to the atoms moving freely for a short while before colliding. In the idealized case, where $\tau_{j1} \ll \tau_{j2} \ll \tau_{j3} \ll \ldots$, one can resolve the different exponents, and in a given interval a single exponent governs the behavior of the correlation function.

A characteristic time constant τ'_j, which in the ideal case of single exponential decay corresponds to the correlation time τ_j, can be estimated from eq 8, where t_a and t_b are two different times.

$$\tau'_j = \frac{t_b - t_a}{\ln\{C_j(t_a)\} - \ln\{C_j(t_b)\}} \tag{8}$$

Time constants τ'_1 for the first-order Legendre polynomial have been estimated from the interval 1–8 ps of the correlation function from eq 8. The time interval was chosen so as to avoid both statistical uncertainties at longer times and initial, rapidly decaying components.

(55) Teleman, O.; Jönsson, B.; Engström, S. *Mol. Phys.*, in press.
(56) Kretsinger, R. H.; Nockolds, C. E. *J. Biol. Chem.* **1973**, *218*, 3313.

Table II. R Factors as a Function of Time (See eq 14) for the Backbone Atom Positions of Parvalbumin Compared to the X-ray Coordinates from Each of the Three Simulations

time after start of simulation, ps	APO	VAC	time after start of simulation, ps	AQ
29	3.05	2.93		
39	3.14	3.01	36	2.16
49	3.21	3.19	46	2.37
59	3.22	3.33	56	2.52
69	3.14	3.36	66	2.43
79	3.27	3.35	76	2.57
89	3.19	3.37	86	2.44
99	3.35	3.46	96	2.53
109	3.32	3.59	106	2.72
mean	3.21	3.29		2.47
RMS deviation	0.09	0.20		0.15

In order to describe the overall rotation of parvalbumin in solution, $C_1(t)$ values were calculated for several long vectors in the protein, and their time constants were estimated by using eq 8. The τ'_1 values for vectors along the helices (cf. below) range from 0.9 (helix A) to 2.5 ns (helix C), and for the vector between the calcium ions $\tau'_1 \approx 1.7$ ns. Since the protein is flexible, it is difficult to define an unambiguous quantity describing the overall rotation. We will therefore use τ'_1 for helix C as an estimate of the overall rotational time constant.

The diffusion coefficient is formally defined as

$$D = \lim_{\Delta t \to \infty} \langle |\mathbf{r}(t + \Delta t) - \mathbf{r}(t)|^2\rangle_t / 6\Delta t \tag{9}$$

where $\mathbf{r}(t)$ is the position of the center of mass and D is the translational diffusion coefficient. The diffusion coefficient can then be calculated from the slope of the mean square displacement vs. Δt in an interval where Δt is large compared to the correlation time (ζ^{-1}) for the velocity autocorrelation function. From the slope in the interval Δt = 20–50 ps, $D = 2 \times 10^{-10}$ m^2/s was obtained. The Einstein relation[57] $D = kT/m\zeta$ gives a correlation time for the velocity correlation function $\zeta^{-1} \approx 1.0$ ps. Thus the interval used is well inside the diffusion limit. The upper limit of the interval is set by the statistical uncertainties due to few points in the averaging.

From a hydrodynamic model, assuming the protein to be spherical, it is possible to obtain an effective radius as well as a viscosity of water from the Stokes–Einstein formula:[58]

$$D = \frac{kT}{6\pi\eta r} \qquad \tau_1 = \frac{4\pi\eta r^3}{kT} \tag{10}$$

Assuming that the Stokes radius is the same for translation and rotation, we obtain

$$r^2 = (1.5 D\tau_1) \tag{11}$$

and

$$\eta^2 = \frac{(kT)^2}{54\pi^2\tau_1 D^3} \tag{12}$$

From the above values of τ'_1 and D we find r = 9.1 Å and $\eta = 1.1 \times 10^{-3}$ Ns/m^2. The calculated radius is to be compared to the radius of 15.5 Å obtained from the radius of gyration. The viscosity is much larger than the approximate value of 0.24×10^{-3} Ns/m^2 obtained for pure water by using the same water model.[59] We believe that the differences in radius and viscosity partly are due to our difficulties in estimating τ'_1 and D because of the shortness of the trajectory. On the other hand, the high viscosity may be correct, since a real 0.02 M protein solution in general is highly viscous.

(57) McQuarrie, D. A. *Statistical Mechanics*; Harper and Row: New York, 1976; p 455.
(58) Marshall, A. G. *Biophysical Chemistry; Principles, Techniques and Applications*; Wiley: New York, 1978; p 712.
(59) Teleman, O.; Ahlström, P. *J. Am. Chem. Soc.* **1986**, *108*, 4333.

Table III. Mean Distances in Å between the Centroids of the Eight Phenylalanines in the Core

residue	Phe 24	Phe 29	Phe 30	Phe 47	Phe 66	Phe 70	Phe 85
Phe 29							
X-ray	5.83						
APO	3.88						
VAC	4.14						
AQ	5.18						
Phe 30							
X-ray	7.75	6.42					
APO	5.17	5.23					
VAC	6.61	6.40					
AQ	5.40	6.11					
Phe 47							
X-ray	16.79	16.23	12.40				
APO	13.06	11.95	10.86				
VAC	19.12	18.86	14.04				
AQ	12.22	15.30	9.58				
Phe 66							
X-ray	6.31	5.70	6.41	11.57			
APO	8.24	6.85	7.81	5.94			
VAC	7.85	8.91	5.22	11.82			
AQ	6.81	9.36	5.26	7.12			
Phe 70							
X-ray	9.44	4.92	8.10	13.80	5.22		
APO	11.29	8.39	10.41	7.58	5.10		
VAC	6.72	6.05	5.21	13.37	4.46		
AQ	8.63	8.30	6.09	10.95	5.10		
Phe 85							
X-ray	5.63	8.75	6.52	12.02	5.36	10.28	
APO	5.93	6.20	5.56	7.42	4.06	8.21	
VAC	5.87	8.06	4.93	16.18	5.34	7.52	
AQ	4.48	9.26	6.59	9.16	5.64	9.76	
Phe 102							
X-ray	11.78	10.97	6.34	6.08	7.36	9.80	7.62
APO	9.80	8.75	6.18	5.25	5.70	7.84	5.49
VAC	13.64	13.42	7.50	7.81	7.73	9.48	10.19
AQ	8.46	11.40	5.75	4.53	5.44	9.22	6.16

General Structure of the Protein. The deformation of the protein molecule was monitored by means of the radius of gyration (r_G), so-called R factors, and changes in distances in the hydrophobic core. The radius of gyration is defined through eq 13, where m_i

$$r_G^2 = \frac{\sum m_i r_i^2}{\sum m_i} \tag{13}$$

is the mass of the ith atom, r_i is its distance to the center of mass of the protein molecule, and the sums run over all atoms in the molecule. For a uniform sphere of radius r, one can show that $r_G = 0.6^{0.5}r$. In the X-ray structure the radius of gyration of parvalbumin is 12.8 Å, corresponding to a sphere radius of 16.5 Å. In all three simulations r_G decreases, but more so in the two in vacuo simulations. At the end of the APO, VAC, and AQ simulations, r_G = 11.5, 11.6, and 12.0 Å, respectively (corresponding to spheres with r = 14.8, 15.0, and 15.5 Å). The smaller radii of gyration in the in vacuo simulations in part reflect that the side chains tend to fold onto the surface to maximize favorable van der Waals interactions. Also, the backbone contracts somewhat in all three simulations, as seen from the distances between all amide nitrogens of the backbone as a function of time.

It has been suggested[60] that proteins should exhibit low-frequency "breathing modes", that is, a periodic contraction and expansion of the whole protein with a frequency of ≈30 cm^{-1}. We have followed the time evolution of the radius of gyration but have not been able to detect any significant periodic motions.

Comparison of the general structure of the protein with the X-ray structure was achieved by means of R factors as applied in the refinement of X-ray crystallography data.[61] The R factor for a configuration was taken to be the minimum

$$R = [\sum_{i=1}^{N} m_i|\mathbf{r}_{i,a} - \mathbf{r}_{i,X}|^2 / \sum_{i=1}^{N} m_i]^{0.5} \tag{14}$$

after rigid body superposition of the backbone with the X-ray conformation; m_i is the mass of atom i and $\mathbf{r}_{i,a}$ and $\mathbf{r}_{i,X}$ its coordinate in the actual configuration and the X-ray structure, respectively. The R factors for the backbone atom positions compared to their positions in the X-ray structure have been calculated every 10 ps for the trajectories (see Table II). The mean R factors from the two in vacuo simulations are about 3.2 Å and from the AQ simulation 2.5 Å. However, in all three cases and especially in the VAC simulation they tend to increase slightly, possibly indicating a continuous change in conformation. Many MD simulations of proteins report root mean square fluctuations around atomic average positions; that is, $\mathbf{r}_{i,X}$ in eq. 14 is replaced by the average coordinate obtained during the simulation. These values, proportional to X-ray temperature factors, are generally in the order of 1 Å, depending on the protein under study as well as the length of simulation. If the protein undergoes large structural changes, then both R factors and "temperature factors" will increase with time. In most simulations this seems to be the case. For example, from a simulation of a BPTI crystal, van Gunsteren et al.[28] report R factors to increase from 1.44 Å after 8 ps to 2.11 Å 11 ps later. Levitt[62] reports R factors of 1.6 and 2.5 Å for C_α atoms and all atoms, respectively, from a 132-ps vacuum simulation of BPTI.

Another method to estimate the differences between the general structure in the three simulations is the distances between different residues. We have chosen to look at the mean distances between the centroids of the phenylalanines situated in the core in the X-ray structure, i.e., all phenylalanines except for Phe 2 and Phe 57 (see Table III). It shows that the structure differs significantly between the different simulations and between each of the simulations and the X-ray structure. The results clearly show that the protein undergoes large changes relative to the initial X-ray

(60) Suezaki, Y.; Go, N. *Int. J. Pept. Protein Res.* **1975**, *7*, 333.

(61) See, for instance: Cantor, C. R.; Schimmel, P. R. *Biophysical Chemistry, Part II: Techniques for the Study of Biological Structure and Function*; Freeman: San Francisco, 1980; p 763–768. The procedure followed here corresponds to eq 13–108 of that ref but in real space and with κ = 1 and W_{hkl} equal to the relative masses of the atoms.

(62) Levitt, M. *J. Mol. Biol.* **1983**, *168*, 595.

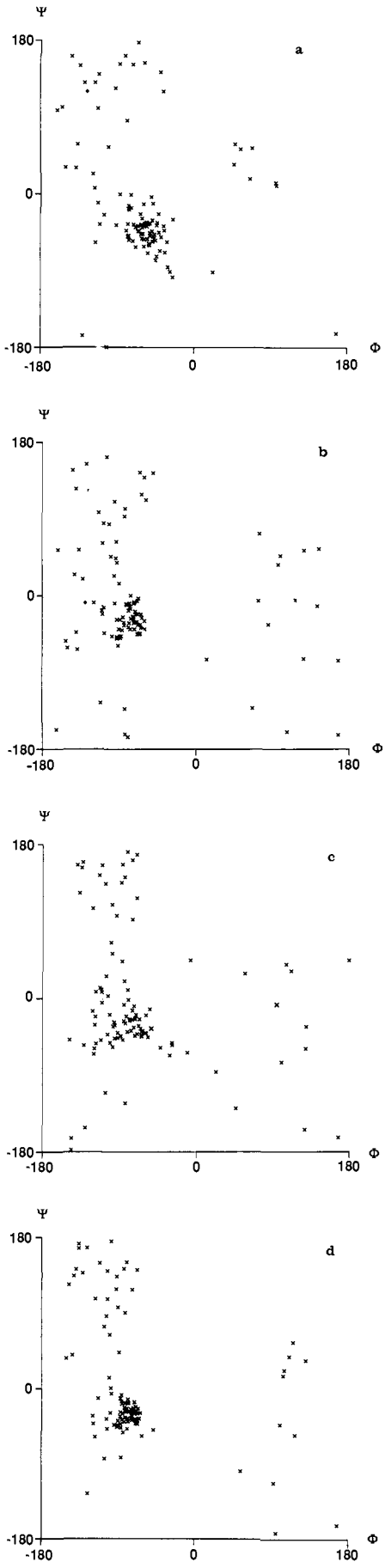

Figure 2. Plot of mean values for backbone dihedral angles Ψ and Φ (a) in the X-ray structure, (b) in the APO simulation, (c) in the VAC simulation, and (d) in the AQ simulation.

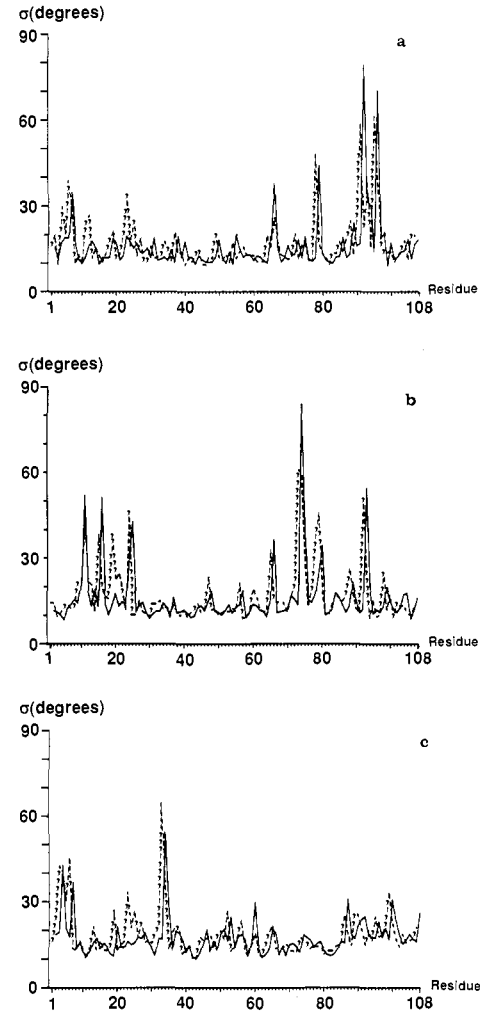

Figure 3. Standard deviations of Ψ (dashed line) and Φ (solid line) along the chain (a) in the APO simulation, (b) in the VAC simulation, and (c) in the AQ simulation.

coordinates. These results and also the R factors reported above may be due to artifacts in the potential function or due to different environments felt by the protein in a crystal, in vacuo (with or without Ca^{2+}) and in solution. This is a question of paramount interest, which unfortunately we cannot answer at present. It is interesting to note the large differences between the in vacuo and the AQ simulations, clearly demonstrating the important role played by the solvent in determining the protein structure. The phenylalanines investigated in Table III are all situated in the core of parvalbumin; hence the structural influence of solvent is not only limited to the exterior part of the protein.

Backbone. It has been suggested that α-helices and β-sheets of a protein act as "springs" and show low-frequency (5–10-cm^{-1}) collective stretching vibrational modes.[63] To investigate these hypotheses, vectors parallel to the helices were defined between atoms at helix endpoints. The lengths of these vectors were calculated as functions of time and Fourier transformed. From the spectrum it was not possible to deduce the presence of slow stretching modes in any of the six helices; i.e., stretching modes seemed to be stochastic in all three simulations. In other in vacuo simulations[64] large-scale vibrational modes have been detected. In most of these the protein under investigation had a domain structure, see, e.g., Åqvist et al.,[64] whereas parvalbumin is a globular protein with no separable domains. However, the solvent considerably affects dynamics all over the protein and may have a damping effect on low-frequency modes. Long-range electro-

(63) Peticolas, W. L. *Methods Enzymol.* **1979**, *61*, 425.
(64) Åqvist, J.; van Gunsteren, W. F.; Leijonmarck, M.; Tapia, O. *J. Mol. Biol.* **1985**, *183*, 461.

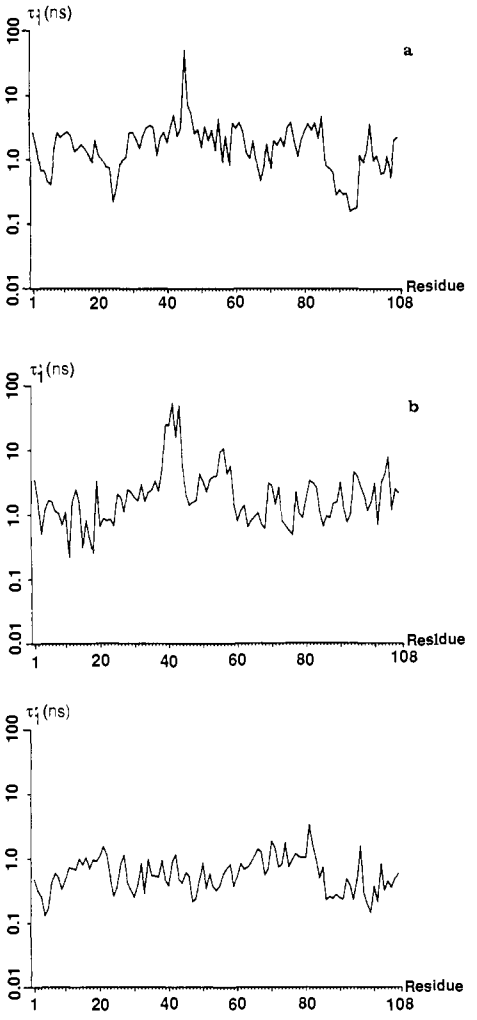

Figure 4. Time constants (τ'_1; eq 8) for all C_α–C_α vectors along the protein backbone (a) in the APO simulation, (b) in the VAC simulation, and (c) in the AQ simulation.

Table IV. Distribution of Hydrogen Bond Lifetimes in the Protein Molecule at Two Formation Distances (d_s, Å) for Each Simulation

	APO		VAC		AQ	
lifetime, ps	2.1	2.3	2.1	2.3	2.1	2.3
<10	413	3137	399	2731	537	3650
10–20	63	186	45	168	63	123
20–30	22	78	28	41	19	25
30–40	7	24	11	29	4	11
40–50	14	11	12	18	6	6
50–60	3	11	2	12	2	3
>60	3	9	8	15	1	2

static interactions, cooperative in character, may also strongly affect the motional behavior of large fragments in a protein.

The structure of a polypeptide chain can be specified by dihedral angles Φ and Ψ. (Φ_i is the angle between the planes C_{i-1}–N_i–$C_{\alpha i}$ and N_i–$C_{\alpha i}$–C_i; Ψ_i is the angle between the planes N_i–$C_{\alpha i}$–C_i and $C_{\alpha i}$–C_i–N_{i+1}.) We calculated the mean of these angles for all residues and plotted Ψ_i against Φ_i for each simulation in Ramachandran maps (Figure 2). In the AQ simulation the points are more clustered than in the VAC and the APO simulations. The density is especially large in an area around $\Phi \approx -70°$ and $\Psi \approx -40°$, close to the α-helix angles ($\Phi = -57°$ and $\Psi = -47°$). In the VAC and APO simulations the spread of points may indicate a beginning denaturation of the α-helices. In all three simulations the area with higher density is moved toward lower Φ and higher Ψ values compared to X-ray data. This change corresponds to a lowering of the number of residues per turn, a change that also can be seen from the hydrogen bond pattern (see below).

The standard deviations were calculated for the dihedral angle

Table V. Calcium Ligands in the VAC and the AQ Simulations[a]

CD site		EF site	
VAC	AQ	VAC	AQ
X-ray Ligands Bound during Whole Simulation			
$O_{\delta 1}$ Asp 51	$O_{\delta 1}$ Asp 51	$O_{\delta 2}$ Asp 90	$O_{\delta 2}$ Asp 90
$O_{\delta 1}$ Asp 53		$O_{\delta 1}$ Asp 92	$O_{\delta 1}$ Asp 92
O_γ Ser 55		$O_{\delta 2}$ Asp 92	
O_{pep} Phe 57		$O_{\delta 1}$ Asp 94	$O_{\delta 1}$ Asp 94
$O_{\epsilon 1}$ Glu 59	$O_{\epsilon 1}$ Glu 59	O_{pep} Lys 96	
$O_{\epsilon 2}$ Glu 62	$O_{\epsilon 2}$ Glu 62	$O_{\epsilon 1}$ Glu 101	
		$O_{\epsilon 2}$ Glue 101	$O_{\epsilon 2}$ Glu 101
X-ray Ligands Bound during Part of Analysis			
	O_{pep} Phe 57 (–eq, 8–)		$O_{\delta 2}$ Asp 92 (–eq, 12–)
	$O_{\delta 1}$ Asp 53 (–2)		O_{pep} Lys 96 (–59)
			$O_{\epsilon 1}$ Glu 101 (–24, 44–)
			O_{water} 12 (–10)
X-ray Ligands Not Bound			
	O_γ Ser 55 (–eq)	O_{water} 12 (o)	
Other Calcium Ligands			
$O_{\delta 2}$ Asp 53 (eq–)	$O_{\delta 2}$ Asp 53 (eq–)	$O_{\delta 1}$ Asp 90 (7–)	$O_{\delta 1}$ Asp 90 (eq–1)
	$O_{\epsilon 2}$ Glu 59 (eq–)		$O_{\delta 2}$ Asp 94 (24–44, *, 58–)
$O_{\epsilon 1}$ Glu 62 (eq–)	$O_{\epsilon 1}$ Glu 62 (eq–)		O_{water} 1973 (3–)
	O_{water} 1987 (eq–)		

[a] The times in picoseconds, after start of the analysis, during which the ligands remain bound are given in parentheses (eq = equilibration period, o = omitted in the simulation, * = distance oscillating between 2.3 and 4.0 Å in the period 44–58 ps).

distributions (Figure 3) in order to investigate the torsional flexibility in different parts of the protein. In most cases the standard deviations due to fluctuations are 10–15°, whereas structural changes are accompanied by larger deviations. In the AQ simulation this occurs near the N-terminal end of the chain and between residues 33 and 34, i.e., at the end of α-helix B, whereas the fluctuations in the α-helices are small. In the VAC simulation large deviations occur at the A helix, between the A and the B helix, and at residues 92 and 93, which are situated near the calcium-loaded EF site. The largest deviations, nearly 90°, occur between the D and E helices. These fluctuations should not be interpreted as a high structural mobility but are due to structural changes taking place during the simulation. In the APO simulation the backbone is very much affected by the absence of calcium, especially near the EF site where several residues show considerable deviations.

The dynamics of the backbone were investigated by means of time correlation functions and their time constants (τ'_1). These were calculated for vectors between α-carbons in neighboring amino acid residues. Typical time constants in the VAC simulation are between 2 and 10 times (in some cases even more) longer than those of the AQ simulation. However, there are also some examples of time constants being longer in the AQ simulation than in the VAC simulation. The results from the APO simulation give τ'_1 values that in general are closer to the values from the VAC simulation.

The largest time constants (in the order of nanoseconds) in the VAC simulation were obtained around the N-terminal end of the C helix. Also, in the APO simulation the largest time constants were obtained in the C helix. In the AQ simulation the largest time constants (about 1 ns, corresponding to half the overall rotational correlation time of the protein) were obtained in the region between the D and E helices. We note that parts with small time constants normally show large standard deviations in the dihedral angles. This means that the time constants quoted in part may reflect structural changes rather than the dynamical

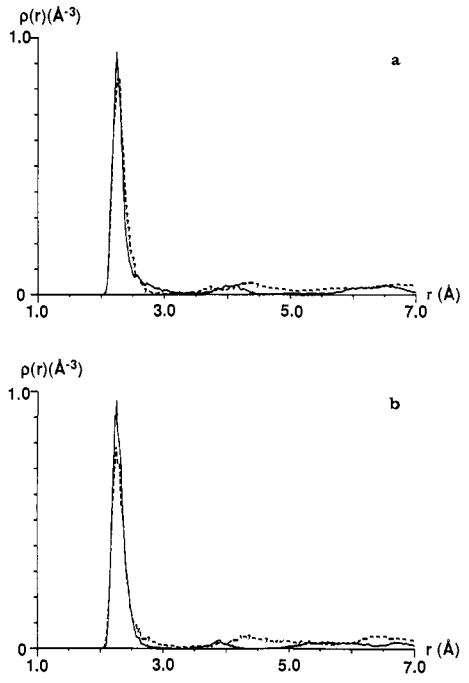

Figure 5. Radial distribution of protein oxygens around the calcium ions for (a) the CD site and (b) the EF site. Solid lines refer to the VAC simulation and dashed lines to the AQ simulation.

Table VI. Time Constants, τ'_1, in ns for Residues in the Calcium Binding Sites

residue	atoms	APO	VAC	AQ
CD site				
Asp 51	C_β–C_γ	1.9	1.7	0.8
Asp 53	C_β–C_γ	1.2	1.9	0.3
Glu 59	C_β–C_δ	0.1	>4.0	0.5
Glu 62	C_β–C_δ	1.3	0.2	0.4
EF site				
Asp 90	C_β–C_γ	0.1	0.8	2.3
Asp 92	C_β–C_γ	0.3	2.4	0.3
Asp 94	C_β–C_γ	0.1	3.2	0.2

behavior. However, we do not think this obscures the interpretation significantly, since we find the most rapid motion in the AQ simulation where the structural changes are least.

Hydrogen Bonds. Since no explicit hydrogen bond potential was used, it is of interest to investigate the behavior of hydrogen bonds in the protein. The definition of a hydrogen bond is not trivial and has to be operational. To investigate possible geometric criteria, Reimers and Watts[65] used a local mode technique to obtain the vibrational spectrum of liquid water. The vibration frequencies were then correlated to a number of geometric properties, such as distances and angles, for intermolecular hydrogen–oxygen pairs constituting potential hydrogen bonds. They found no simple and stringent geometric condition for the formation or existence of hydrogen bonds. Below, we will interpret simulation geometries in terms of hydrogen bonds, but we are aware that this designation is to be regarded with some caution.

A list of all polar hydrogens (i.e., those hydrogens that are taken into account explicitly in the simulation) and a list of all possible acceptors of hydrogen bonds (i.e., polar oxygens and nitrogens) were constructed. Water hydrogen bonds were not considered in this analysis. The distances between all possible combinations were calculated throughout the period of analysis. We chose to use an asymmetric criterion for hydrogen bonds such that a bond is considered to be formed when the distance between hydrogen and acceptor becomes less than some distance d_s and to cease when the distance becomes larger than d_e, where $d_s < d_e$. The reason for this is that encounters at long distance may not form any

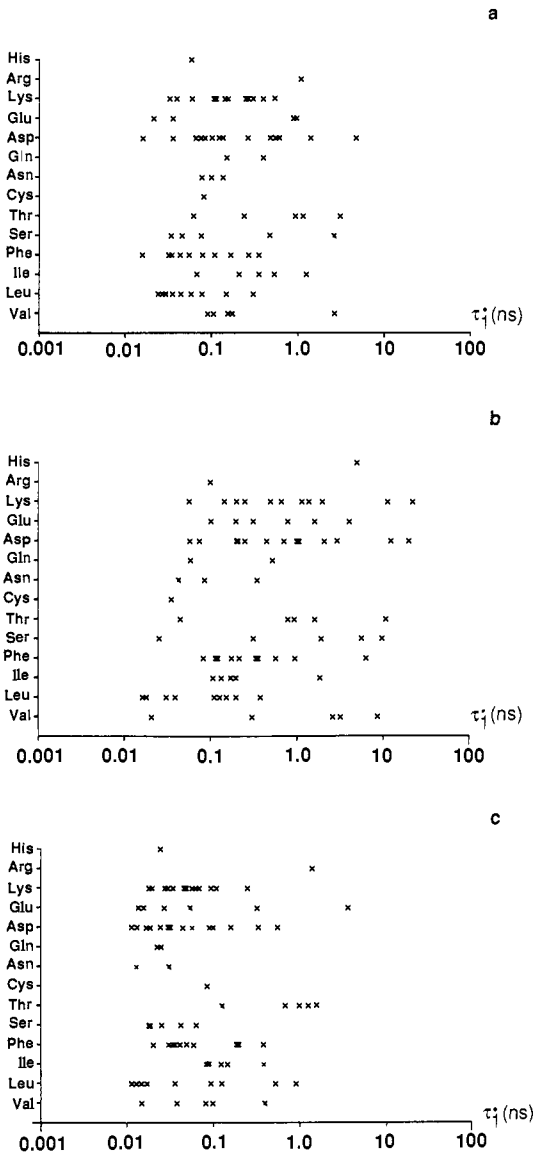

Figure 6. Time constants (τ'_1) for the vector between the two outermost non-hydrogen atoms in all side chains ordered after residue type (a) in the APO simulation, (b) in the VAC simulation, and (c) in the AQ simulation.

hydrogen bond. On the other hand, a hydrogen bond already formed may vibrate and thus occasionally be very long without actually breaking. With d_s = 2.1 and 2.3 Å (d_e = 3.5 Å in both cases) we obtained the results shown in Table IV. We note that hydrogen bonds in the protein are rather short lived—especially in aqueous solution—and most hydrogen bonds lasted less than 10 ps. This is consistent with the properties of pure water. At room temperature it takes a water molecule ≤10 ps to diffuse over a distance comparable to a molecular diameter. This can be taken as an upper limit for the hydrogen bond lifetime. It is assumed here that the diffusing species in water is a single water molecule. Results similar to ours have been obtained by Levitt,[66] who studied the hydrogen bonds in BPTI during a 56-ps in vacuo simulation. He obtained large fluctuations in bond lengths for most of the hydrogen bonds in the α-helices between a peptide oxygen and the peptide hydrogen four residues away. These fluctuations would in our description correspond to several ruptures and reformations of the same hydrogen bond, thus giving short lifetimes for hydrogen bonds in the α-helices.

It is notable that peptide oxygens in α-helices switch between being hydrogen bonded to the peptide hydrogen three residues away and the one four residues away. This is contrary to the

(65) Reimers, J. R.; Watts, R. O. *Chem. Phys.* **1984**, *91*, 201.

(66) Levitt, M. *Nature (London)* **1981**, *294*, 379.

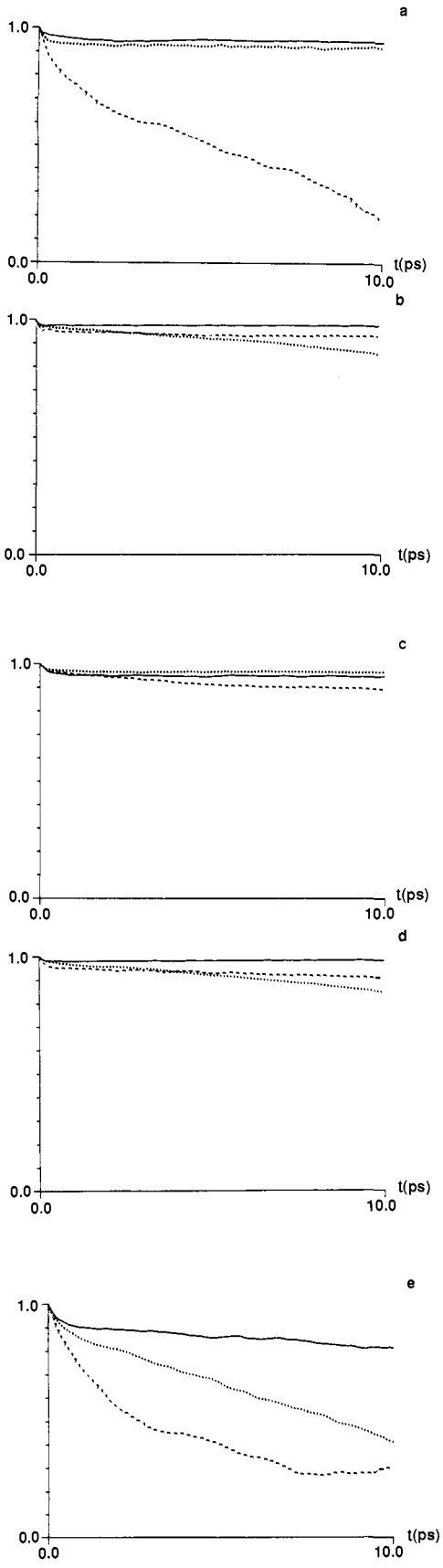

Figure 7. Time correlation functions for vectors in some residues. Solid lines refer to the VAC, dashed lines to the AQ, and dotted lines to the APO simulation. (a) $O_{\delta 1}$–$O_{\delta 2}$ in Asp 8. This residue is situated at the surface of the protein. (b) C_β–$C_{\gamma 1}$ in Val 43, which is situated at the end of the C helix opposite the CD site. (c) $C_{\gamma 2}$–$C_{\delta 1}$ in Ile 50. It is directed into the core and is situated near the CD site. (d) C_β–C_γ in Asp 94, a calcium ligand in the EF site. (e) $O_{\epsilon 1}$–$O_{\epsilon 2}$ in Glu 60, which is situated in the CD site, but does not ligand the calcium ion.

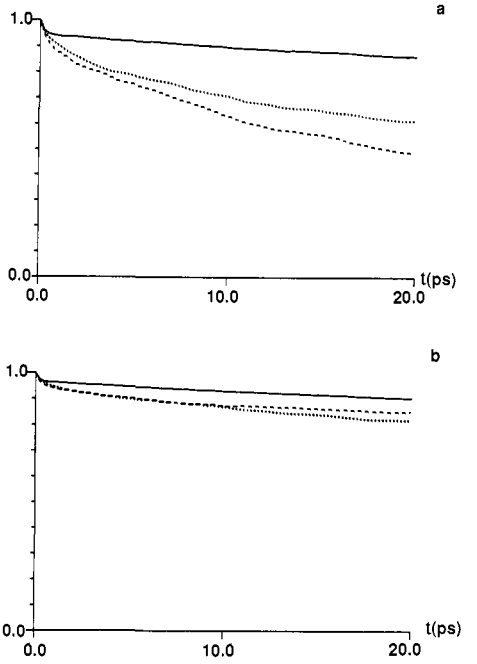

Figure 8. Sum over all phenylalanines of time correlation functions ($C_1(t)$) for (a) DD vectors and (b) GZ vectors. Solid lines refer to the VAC, dashed lines to the AQ, and dotted lines to the APO simulation.

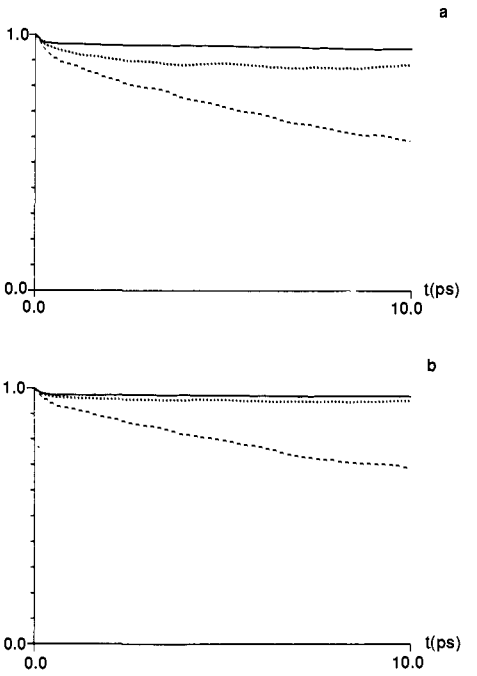

Figure 9. Time correlation functions ($C_1(t)$) for vectors in Phe 30 for (a) DD vectors and (b) GZ vectors. Solid lines refer to the VAC, dashed lines to the AQ, and dotted lines to the APO simulation.

classical picture, based on among others hydrogen-exchange studies, of helices as static. However, there exists no unambiguous relationship between hydrogen bond lifetimes and hydrogen-exchange rates since the latter, as Levitt points out, depend on several factors other than hydrogen bond strength. Our results are also supported by recent results from neutron diffraction studies on ribonuclease,[67] where alternating hydrogen bonds have been observed between the ith peptide oxygen and the $i + 3$ as well as the $i + 4$ peptide hydrogen.

Calcium Binding Sites. In the X-ray structure, the calcium ion in the CD site is liganded to six protein ligands and that in the EF site to seven protein ligands and one crystal water. In order

(67) Sjölin, L., personal communication.

to characterize calcium ligandation we calculated the distances between the calcium ions and all possible ligands. These were regarded to be ligands if closer to the calcium ion than 3.0 Å. The results from this analysis are shown in Table V.

The crystal water from the X-ray structure was not included in the in vacuo simulations. No protein ligand was released during the VAC simulation, but further carboxyl oxygens became liganded. In the CD site the calcium acquired two new ligands during the equilibration, $O_{\delta 2}$ from Asp 53 and $O_{\epsilon 1}$ from Glu 62. At the EF site $O_{\delta 1}$ from Asp 90 arrives after 7 ps of analysis, which means that both binding sites coordinate eight ligands. These three new ligands all belong to carboxyl groups where the other oxygen atom was already a ligand; thus there are no drastic changes in the ligandation. Similar results have been seen in simulations of CaEDTA complexes.[59] The radial distribution functions for both sites show sharp peaks at 2.3 Å (Figure 5), which is in good agreement with X-ray data on calcium chelation in proteins.[68] The ligands are not very mobile, and the time constants τ'_1 for vectors in the liganding residues are with a few exceptions in the order of nanoseconds (see Table VI).

In the APO simulation, of course, the picture is quite another. In the absence of calcium ions, mobility is enhanced and structure is lost. The only exceptions are a few ligands in the CD site which are about as immobile as those in the VAC simulation (see Table VI). This is probably due to a neighboring positively charged lysine (Lys 54), which stabilizes the negatively charged ligands.

The results from the AQ simulation differ in many respects from those of the VAC simulation. Crystal waters are present from the beginning, and ligands exchange with water molecules from the solution. A water molecule enters the EF site after a few picoseconds of analysis and remains. The crystal water leaves the EF site after 10 ps. In the CD site a water molecule enters during the equilibration and remains. The protein ligands are less tightly bound to the calcium in water than in vacuo. Several protein ligands leave the calcium ion for a short while but return again. So is $O_{\epsilon 1}$ in Glu 101 released after 24 ps of the analyzed trajectory but returns 20 ps later. The calcium ions loose one protein ligand each compared to the X-ray structure. In the CD site, O_γ of Ser 55 departs after less than 10 ps of equilibration, and in the EF site, the peptide oxygen of Lys 96 leaves after 59 ps of analysis. Further changes of ligands are frequent and are specified in Table V.

An interesting result is that coordination numbers are equal (≈8) at both calcium binding sites during most of the AQ and VAC simulations, whereas they are 6 at the CD site and 8 at the EF site in the X-ray structure. Other MD simulations give coordination numbers of ≈7 or ≈9 for a calcium ion in water[59,69] and ≈6.7 for a calcium ion in a $CaEDTA^{2-}$ complex.[59]

The dynamics of protein ligands differ between the three simulations. This can be seen from the τ'_1 values for vectors between C_β and the carboxyl carbon in the nonexchanging ligands. These values are shown in Table VI (see also Figure 7). From the table it is clear that mobility in general is enhanced in the AQ simulation compared to the VAC simulation.

Side Chains. Time correlation functions, $C_1(t)$, and corresponding time constants from the inverval 1–8 ps were computed for the vector between the two outermost non-hydrogen atoms in all side chains except for alanines and glycines (Figure 6). From a comparison of results from the two in vacuo simulations, one notices changes in time constants all over the protein upon removal of the calcium ions. Mostly, the mobility is enhanced, but in some cases there are changes going in the opposite direction.

A tendency valid for all types of residues is that time constants are shorter in aqueous solution than in vacuo. This is very pronounced for charged and polar side chains at the surface, such as glutamic and aspartic acids, lysines, and serines (see also Figure 7) and nearly as pronounced for the hydrophobic phenylalanines in the core of the protein. Therefore we have analyzed the behavior of phenylalanines more carefully.

(68) Einspahr, H.; Buggi, C. E. *Met. Ions Biol. Syst.* **1983**, *17*, 51.
(69) Probst, M. M.; Radnai, T.; Heinzinger, K.; Bopp, P.; Rode, B. M. *J. Phys. Chem.* **1985**, *89*, 753.

Table VII. Minimum Distances between Phe Groups and Water Oxygens during the Simulation and Time Constants, τ'_1, in ps for Vectors in the Phe Groups

residue number	minimum distance, Å	τ'_1 APO GZ	APO DD	VAC GZ	VAC DD	AQ GZ	AQ DD
2	2.8	21	5	>4000	>4000	147	8
24	4.7	343	214	106	113	40	28
29	3.2	299	304	358	247	428	177
30	3.9 (7)[a]	380	107	966	374	41	21
47	2.9	260	21	70	62	170	19
57	2.8	34	52	2000	927	70	19
66	6.6	108	65	1671	397	321	42
70	4.4	119	17	1150	193	733	20
85	2.9 (6)[a]	355	35	160	172	146	117
102	2.9 (5)[a]	458	85	221	128	248	303

[a] One water molecule enters the hydrophobic core for a few picoseconds and gives rise to irrelevant shortest distances between some of the phenylalanines and the nearest water. The minimum distances to the second nearest waters are given in parentheses for those phenylalanines.

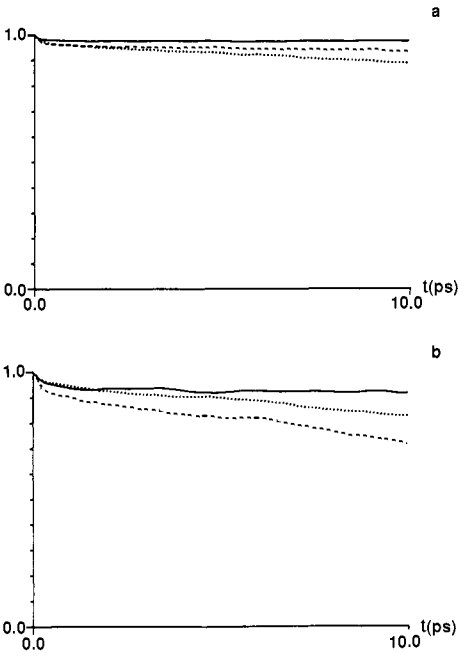

Figure 10. Time correlation functions ($C_1(t)$) for vectors in Phe 66 for (a) DD vectors and (b) GZ vectors. Solid lines refer to the VAC, dashed lines to the AQ, and dotted lines to the APO simulation.

Time correlation functions have been calculated for the vector between $C_{\delta 1}$ and $C_{\delta 2}$ (transverse to the ring bond axis), denoted DD, and the vector between C_γ and C_ζ (approximately parallel to the ring bond axis), denoted GZ. The sums of these correlation functions are shown in Figure 8. It is clear that the effect of water is large—about as large as the major changes induced by the removal of calcium ions. We also calculated time constants as described above and obtained for the DD vector $\tau'_1 \approx 190$ ps in the VAC and 28 ps in the AQ simulation and for the GZ vector 250 and 109 ps, respectively. One could argue that the difference between the simulations is due merely to those phenylalanines situated at the surface of the protein. To clarify this, we also plotted the individual time correlation functions and tabulated their time constants (Table VII). Two examples, Phe 30 and Phe 66, both situated in the core of the protein, are given in Figures 9 and 10. In the AQ simulation there is a rapid decay of the correlation function in the beginning. This is probably due to the protein atoms being less densely packed than in vacuo and hence having a longer mean free path. Another feature is the large difference between DD and GZ time constants. This implies a considerable rotational motion of the ring around the C_γ–C_ζ axis. This can also be seen from the angle between the plane defined by C_α, C_β, and C_γ and that by C_γ, $C_{\delta 1}$, and $C_{\delta 2}$. It shows fast

rotational motions spanning about 30° and some occasional larger turns. Phe 30 and Phe 66 both have increased mobility in water compared to in vacuo, and DD time constants are 374 ps for Phe 30 and 397 ps for Phe 66 in the VAC simulation and 20 and 42 ps, respectively, in the AQ simulation. Similar results are obtained for several other internal phenylalanine residues as well, which shows that the presence of water does indeed influence the dynamics not only of external parts of the protein but also of the interior.

Conclusion

The protein structure deviates from its crystal form in all three simulations. This is, of course, a consequence of the potential, but whether it is merely an artifact or a representation of the real conformation remains to be shown. It is certain, though, that the parvalbumin molecule has to be more or less flexible in order to function. Simulation confirms this picture of a fairly flexible molecule.

Water is paramount, and its presence affects the dynamics and structure of the entire protein. Large effects on dynamics are found not only at the surface but also in the interior. The protein structure in aqueous solution more resembles the crystal forms than do the in vacuo ones, as manifested by surface side chains and backbone dihedral angles.

Ligandation of the calcium ions is slightly different for the two binding sites in the crystal form, but not so in the simulations. Exchanges of water ligands do occur, which implies that binding sites are solvent accessible as, indeed, they have to be. Removal of the calcium ions affects local dynamics in all parts of the molecule. The protein is thus capable of fast, global information transfer, a prerequisite of the strong calcium binding cooperativity found from experiment.[70]

The overall translational and rotational diffusions of the protein molecule were each obtained within the correct order of magnitude despite the short duration of the simulation. This is due to the fact that the statistics of these properties are determined by the solvent more than by the solute.

These simulations, especially that in aqueous solution, present a molecule with several interesting properties consistent with the function of parvalbumin. Thus, molecular dynamics simulation is a complement to experiment, even for these very complex systems. For simulation results to be more reliable in detail, though, much validation of inter- and intramolecular potentials as well as more technical aspects remains. In view of the far-reaching advances in the understanding of biomolecular mechanisms then possible, however, these issues deserve great effort.

Acknowledgment is due to Dr. Torbjörn Drakenberg for valuable discussion and encouragement and to the Swedish Natural Science Research Council for generous allocation of computer time.

Registry No. Ca, 7440-70-2.

Supplementary Material Available: Tables of atom types and charges, Lennard-Jones potential parameters, bond length and angle potential parameters, and dihedral angle potential parameters (13 pages). Ordering information is given on any current masthead page.

(70) Teleman, O.; Drakenberg, T.; Forsén, S.; Thulin, E. *Eur. J. Biochem.* **1983**, *134*, 453.

Vibrational Circular Dichroism in Transition-Metal Complexes. 3. Ring Currents and Ring Conformations of Amino Acid Ligands

Teresa B. Freedman,* Daryl A. Young, M. Reza Oboodi, and Laurence A. Nafie*

Contribution from the Department of Chemistry, Syracuse University, Syracuse, New York 13244-1200. *Received August 5, 1986*

Abstract: The enhanced NH and CH stretching vibrational circular dichroism (VCD) spectra of tris(amino acidato)cobalt(III) complexes in acidic aqueous solution and bis(amino acidato)copper(II) complexes in Me_2SO-d_6 solution are interpreted in terms of vibrationally generated ring currents. In the complexes, hydrogen-bonding interactions between an NH bond of one ligand and either an oxygen lone pair (meridional and planar complexes) or a carbonyl π orbital (facial complexes) on an adjacent ligand are identified which restrict the ligand conformations in solution. The VCD spectra are most consistent with the ligand envelope conformations puckered at the metal atom. The large biased NH stretching VCD arises from oscillating current in the ligand and hydrogen-bonded rings generated by the NH vibrational motion.

Vibrational circular dichroism (VCD)[1-5] provides a unique probe for the conformation of chiral molecules in solution, since each vibrational normal mode corresponds to a specific and often localized distortion of the nuclear framework. VCD arises from the generation of nonorthogonal electric and magnetic dipole moments during vibrational excitation. The sign and magnitude of the VCD intensity depend on the manner in which the electrons respond to the nuclear motion. When the electronic motion perfectly follows the nuclear motion, bisignate VCD features may be observed, which are adequately described by the coupled oscillator (CO)[5] or fixed partial charge (FPC)[6] intensity models.

(1) Keiderling, T. A. *Appl. Spectrosc. Rev.* **1981**, *17*, 189.
(2) Nafie, L. A. In *Advances in Infrared and Raman Spectroscopy*; Clark, R. J. M., Hester, R. E., Eds.; Wiley-Heyden: London, 1984; Vol. II, p 49.
(3) Stephens, P. J.; Lowe, M. A. *Annu. Rev. Phys. Chem.* **1985**, *36*, 213.
(4) Freedman, T. B.; Nafie, L. A. In *Tropics in Stereochemistry*; Eliel, E., Wilen, S., Allinger, N., Eds.; Wiley: New York, 1986; Vol. 17, in press.
(5) (a) Holzwarth, G.; Chabay, I. *J. Chem. Phys.* **1972**, *57*, 1632. (b) Sugeta, H.; Marcott, C.; Faulkner, T. R.; Overend, J.; Moscowitz, A. *Chem. Phys. Lett.* **1976**, *40*, 397.
(6) (a) Schellman, J. A. *J. Chem. Phys.* **1973**, *58*, 2282; **1974**, *60*, 343. (b) Deutsche, C. W.; Moscowitz, A. *J. Chem. Phys.* **1968**, *49*, 3257; **1970**, *53*, 2630.

0002-7863/87/1509-1551$01.50/0