

# Anharmonic Wave Functions of Proteins: Quantum Self-Consistent Field Calculations of BPTI

Adrian Roitberg, R. Benny Gerber, Ron Elber, Mark A. Ratner\*

The harmonic approximation for the potential energy of proteins is known to be inadequate for the calculation of many protein properties. To study the effect of anharmonic terms on protein vibrations, the anharmonic wave functions for the ground state and low-lying excited states of the bovine pancreatic trypsin inhibitor (BPTI) were calculated. The results suggest that anharmonic treatments are essential for protein vibrational spectroscopy. The calculation uses the vibrational self-consistent field approximation, which includes anharmonicity and interaction among modes in a mean-field sense. Properties obtained include the quantum coordinate fluctuations, zero-point energies, and the vibrational absorption spectrum.

Quantitative understanding of the physical properties of proteins at low temperatures inherently requires quantum mechanical treatment. For instance, quantum calculations are essential for the quantitative analysis of vibrational spectroscopy; low-temperature Debye-Waller attenuation factors of x-ray scattering intensities are subject to large quantum effects; the issue of tunneling states in proteins is a purely quantum mechanical area; and zero-point energy effects are expected to play a major role in protein energetics and in the interaction of proteins with water molecules.

Quantum treatment of protein vibrations was hitherto carried out mostly at the level of harmonic normal mode treatments. However, the many large-amplitude, low-frequency modes of proteins, the "softness" of available protein potential functions for many types of displacements, and the many nearly equivalent minima of these potentials (1, 2) all suggest the inadequacy of the harmonic approximation for proteins. Efforts to include some anharmonic effects in a quantum framework (3) did not include quantities such as wave functions, which are essential for spectroscopy.

In this report, we put forward a quantitatively adequate, practically feasible approach for computing the ground and the low-lying vibrational eigenfunctions of proteins. We use the vibrational self-consistent field (SCF) method, which includes the effects of anharmonic interactions between the different normal modes in an approxi-

mate framework (4), to analyze a model potential of a solvated protein, the bovine pancreatic trypsin inhibitor (BPTI). We present arguments why this approximation should be reasonably good for the ground and low-lying excited states considered here. The calculated ground-state and excited-state vibrational wave functions of the system, and the resulting vibrational absorption spectrum, illustrate the feasibility of quantitative theoretical treatment of protein properties. We also calculate several other properties, such as the zero-point energies and the amplitudes (or position fluctuations) of the various modes, to assess the importance of the quantum effects. Finally, we compare our calculations with results in the framework of the harmonic normal mode model, to throw light on the limitations of this important approximation.

The BPTI protein inhibits trypsin by tightly binding to it at Lys<sup>15</sup>. It contains 58 amino acids with three disulfide bonds in the native form. It has been extensively studied and has been established as a regular test case to which new methods for understanding protein structure and dynamics are applied. In particular, normal mode analysis of BPTI has been carried out by several groups (5–10), and neutron scattering has been used to examine its very low frequency vibrational motions (5, 11).

The dynamics and potential energy of the system were studied with the program MOIL (12, 13). Although there is no guarantee that this potential is sufficiently accurate to predict spectra, it does provide an attractive standard potential model for testing the extent to which quantum and anharmonic effects in the potential change the observable properties of the molecule. The total number of "atoms" in the system was 1147, corresponding to 3441 degrees of freedom, or a 3435-dimensional wave function when the rigid-body constraints are included.

We treated anharmonicities by expanding the Hamiltonian to fourth order in the normal coordinate set (14). This gives

$$H = \sum_{k=1}^{3N-6} T_k + \frac{1}{2} \sum_{k=1}^{3N-6} V_{kk}'' Q_k^2 + \frac{1}{6} \sum_{k=1}^{3N-6} V_{kkk}^{(3)} Q_k^3 + \frac{1}{24} \sum_{k=1}^{3N-6} V_{kkkk}^{(4)} Q_k^4 + \frac{1}{2} \sum_{k=1}^{3N-6} \sum_{l \neq k}^{3N-6} V_{kkl}^{(3)} Q_k^2 Q_l + \frac{1}{4} \sum_{k=1}^{3N-6} \sum_{l \neq k}^{3N-6} V_{kkll}^{(4)} Q_k^2 Q_l^2 \quad (1)$$

The entire Hamiltonian is expressed in terms of the normal coordinates  $Q_k$ , which were calculated by direct diagonalization of the full mass-weighted quadratic potential expressed in cartesian coordinates. The terms  $T_k$ ,  $V_{kk}''$ , and  $V_{kkk}^{(3)}$  and  $V_{kkll}^{(4)}$  represent, respectively, the kinetic energy in the  $k$ th mode, the force constant for the  $k$ th mode, and cubic off-diagonal anharmonicity coefficients. The  $V_{kkk}''$ ,  $V_{kkl}^{(3)}$ , and  $V_{kkll}^{(4)}$  are partial derivatives of the full potential, evaluated at the minimum. The required third and fourth derivatives were calculated as finite differences from the second derivative matrix; this calculation proved to be numerically stable.

The quartic approximation for the potential should be adequate here because our focus is on the vibrational ground state and the lowest excitations (single quantum) only. For these low-lying states, the quartic approximation is probably valid even for low-frequency torsional modes. The quartic level is expected to break down for higher excitations and is not valid for proteins at room temperature. Support for the validity of the quartic expansion of low-frequency modes was obtained by exploration of the exact energy surface along these coordinates. A favorable comparison to the quartic expansion was obtained.

The SCF technique retains the normal coordinate notion of mode separability, while including diagonal anharmonic effects exactly and off-diagonal effects in a mean-field sense. In the vibrational SCF model, the Hamiltonian is expressed in terms of some independent mode Hamiltonians, and the overall wave function is then a simple Hartree product of single-mode wave functions. That is, we take for the  $n$ -mode wave function the trial form  $\Psi(Q_1 \dots Q_n) = \Phi_1(Q_1) \cdot \Phi_2(Q_2) \cdot \Phi_3(Q_3) \dots \Phi_n(Q_n)$ . With these single-mode wave functions, the effective single-particle Hamiltonians are self-consistently redetermined until convergence is reached. The method is described extensively elsewhere (3, 15–17). Vibrational SCF was used for quartic potentials in the case of formaldehyde (18).

When the SCF procedure is applied to

A. Roitberg and M. A. Ratner, Department of Chemistry, Northwestern University, Evanston, IL 60208, USA.

R. B. Gerber, Department of Physical Chemistry and Fritz Haber Institute, Hebrew University, Jerusalem 91904, Israel, and Department of Chemistry, University of California, Irvine, CA 92717, USA.

R. Elber, Department of Physical Chemistry and Fritz Haber Institute, Hebrew University, Jerusalem 91904, Israel, and Department of Chemistry, University of Illinois, Chicago, IL 60680, USA.

\*To whom correspondence should be addressed.

the Hamiltonian, we obtain for mode  $k$  the following expression for the mean-field potential energy

$$V_k(Q_k) = \frac{1}{2} V_{kk}'' Q_k^2 + \frac{1}{6} V_{kkk}^{(3)} Q_k^3 + \frac{1}{24} V_{kkkk}^{(4)} Q_k^4 + \frac{1}{2} \sum_{l \neq k}^{3N-6} V_{kkl}^{(3)} \langle Q_l \rangle Q_k^2 + \frac{1}{2} \sum_{l \neq k}^{3N-6} V_{kll}^{(3)} \langle Q_l^2 \rangle Q_k + \frac{1}{4} \sum_{l \neq k}^{3N-6} V_{kkll}^{(4)} \langle Q_l^2 \rangle Q_k^2 \quad (2)$$

Here, the averaged values  $\langle Q_l \rangle$  and  $\langle Q_l^2 \rangle$  are evaluated over the single-mode eigenfunction  $\Phi_l(Q_l)$ , which satisfies the single-mode SCF equation

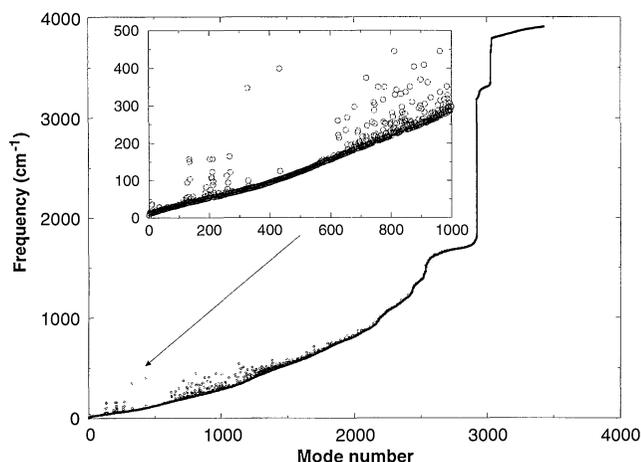
$$(T_l + V_l - \epsilon_l) \Phi_l = 0 \quad (3)$$

Three sets of calculations were performed. In the first, only the harmonic normal modes were used. The second included the diagonal anharmonic cubic and quartic terms. The third involved diagonal as well as off-diagonal terms with the SCF treatment averaged over the lowest 88 modes (19).

Figure 1 shows the effect of diagonal anharmonicity for all modes. For high frequencies ( $>1000 \text{ cm}^{-1}$ ), anharmonic effects are small; this is not surprising because these motions correspond to covalent stretches and angle bends, which are harmonic within the potential energy surface used. Medium to large diagonal anharmonic effects are observed in the low-frequency range. Some of the frequencies changed by more than a factor of 4. Note that all diagonal anharmonic corrections are positive.

Figure 2 shows the normal coordinate frequency for the 88 lowest frequency modes and those including the diagonal and off-diagonal anharmonicity corrections. The squares correspond to exact treatment of the diagonal anharmonicities and SCF treatment of the off-diagonal ones; they almost perfectly correspond to the circles, which only include the diagonal anharmonicities. Thus, the inclusion in the SCF treatment of off-diagonal anharmonicities has almost no effect on the calculated frequency dispersion curve. This is in contrast to small molecule studies, when off-diagonal elements have a strong effect on the frequency dispersion curve (17). One reason for this substantial difference is that many modes enter into the SCF potential; in this 88-mode calculation, for example, each Hamiltonian includes an average over 87 other modes. Because these mixing terms can be either positive or negative, the correction contributions tend to cancel as the number of modes gets very large; this behavior is qualitatively different from that found in small molecules.

The anharmonic terms also affect atom-



**Fig. 1.** The effect of diagonal anharmonicity on all modes. The line represents the harmonic normal mode frequencies as a function of the normal mode coordinate number, and the circles show the corresponding frequencies, including cubic and quartic diagonal anharmonicities. **(Insert)** The same data for the modes with the lowest 1000 frequencies.

ic mean square displacement,  $\text{MSD}^{(i)} = \langle \Psi_{\text{GS}} | r^2(i) | \Psi_{\text{GS}} \rangle$  (Fig. 3). The MSD is related to the experimental Debye-Waller factor of the  $i$ th particle; here  $\Psi_{\text{GS}}$  is the ground-state wave function. Because we have used the united-atom approximation, this value is shown for the heavy atoms only, and hydrogens are not shown explicitly. Hydrogens cannot be detected in x-ray data, the most frequent source of Debye-Waller factors.

The MSD at the normal coordinate level is shown in Fig. 3A. The atoms with sequence numbers greater than 560, corresponding to  $\text{H}_2\text{O}$  molecules, have higher attenuation factors than the protein itself because their MSDs are much larger (they are not held together by covalent forces).

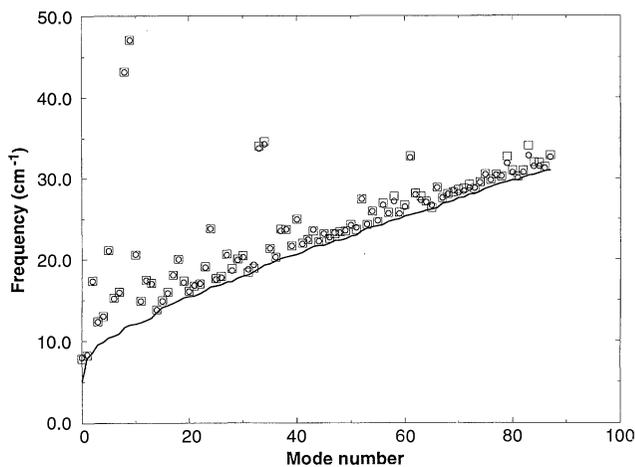
Although the diagonal anharmonic MSD and that from the harmonic normal coordinate analysis is similar for most atoms (Fig. 3B), there are some substantial differences. The two largest differences, below atom number 400, are attributable to two particular residues, Phe<sup>22</sup> and Phe<sup>33</sup>. They undergo large torsional fluctuations, corresponding to very low frequency modes, and therefore show large peaks in the MSD.

Here the correction due to diagonal anharmonicity is substantial.

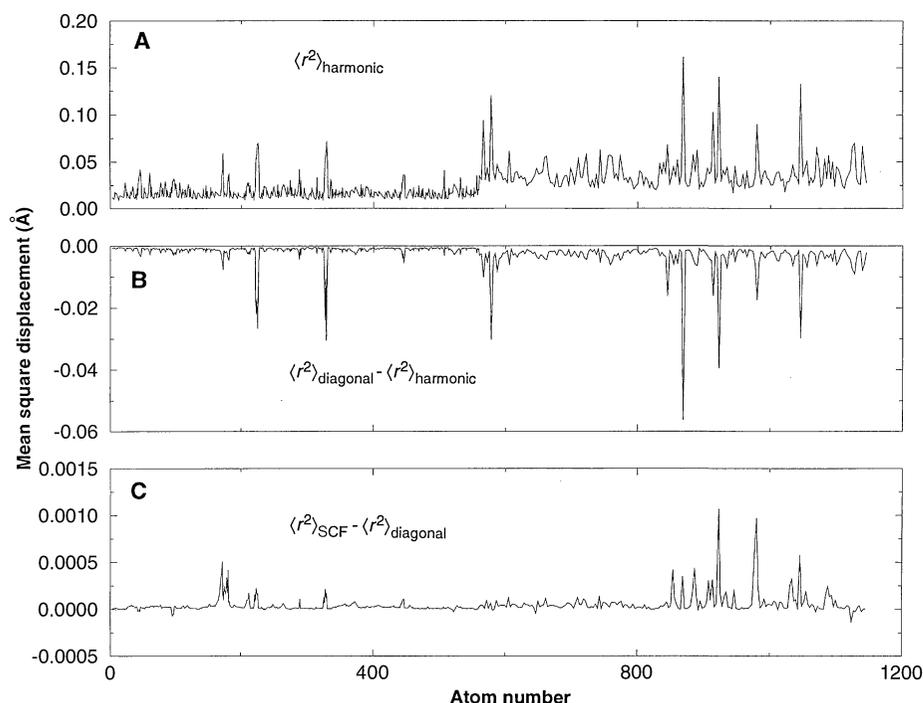
Figure 3C shows the difference between the SCF results with the off-diagonal terms and the diagonal anharmonic MSDs. The net effect of mode mixing, at the SCF level, is relatively small on the observable MSD, similar to the observations in Fig. 2. Notice that the fluctuations seem to average about zero, in agreement with our previous argument about the reason for the relatively small importance of SCF corrections.

Figure 4 shows the absolute difference between the diagonal anharmonic and purely harmonic MSDs, calculated as a function of the harmonic MSD. As expected, the more an atom moves, the more likely it is that the diagonal anharmonic correction will be large, as it is more likely to hit a repulsive wall because of its large zero-point oscillations. Again, there are two very large contributions from Phe<sup>22</sup> and Phe<sup>33</sup>.

Finally, the calculation of the infrared intensity  $I$  as a function of the mode frequency  $\nu_k$  (Fig. 5) was performed with the formula  $I(\nu_k) = \nu_k(d_x^2 + d_y^2 + d_z^2)$ . The transition dipole moments  $d_x$ ,  $d_y$ , and  $d_z$  are calculated by expansion in normal-mode



**Fig. 2.** The effect of off-diagonal anharmonicities as SCF results for the lowest 88 modes. The line corresponds to the harmonic normal mode calculation, the circles include the diagonal terms, and the squares are the SCF results.

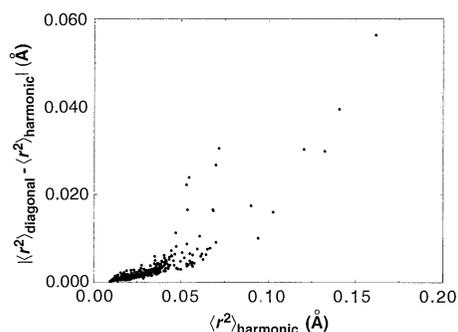


**Fig. 3.** Mean square displacements averaged over the ground-state wave function as a function of atom number. (A) Harmonic normal mode calculation. (B) Difference between the anharmonic diagonal results and those from the harmonic normal mode calculation. (C) Difference between anharmonic SCF and anharmonic diagonal MSDs.

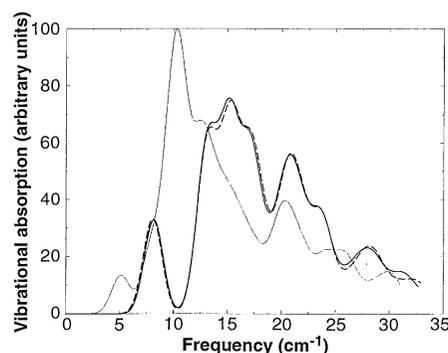
transition moments  $\langle \Phi_k^{(0)} | Q_k | \Phi_k^{(1)} \rangle$ . The results are broadened with Gaussians that have a full width at half maximum of  $2 \text{ cm}^{-1}$ . Again, the results of the diagonal anharmonic computation are substantially different from those of the harmonic normal coordinate calculation, and the SCF approximation makes no significant additional change.

The potential fields available to date for proteins are an important accomplishment for the discipline and have made possible the development of molecular dynamics, Monte Carlo, and other theoretical studies; however, the concern remains that these potentials may be flawed, perhaps seriously

so. We believe that the ability to carry out theoretical quantum calculations on protein vibrations will make it possible to use data from spectroscopy on protein systems to refine potentials. Such refinements have been important for small-molecule systems (20). Although there are great obstacles to extensive vibrational spectroscopic studies of proteins, we believe that there are substantial possibilities for experimental development. Vibrational absorption spectroscopy, Raman spectroscopy, and inelastic neutron scattering relate directly to the calculations and to the theoretical capabilities described here. Comparison of predicted



**Fig. 4.** The absolute value of the difference between the diagonal anharmonic MSDs and the harmonic MSDs (Fig. 3B) as a function of the harmonic MSD (Fig. 3A).



**Fig. 5.** The low-frequency vibrational spectrum for BPTI. The thin line represents the harmonic results, the thick line includes diagonal anharmonic effects, and the dashed line shows the SCF results.

and observed spectra may well be able to markedly improve existing force fields (21).

The quantum effects found here are substantial and suggest that it is essential to include them outside the context of spectroscopy. This implies major effects on Debye-Waller factors and other properties of interest. The effects of zero-point energy, for example, on water adhesion on proteins merit attention in a quantum framework.

Finally, in addition to their biological importance, proteins are glassy-like materials, and such materials are known to have unusual low-temperature properties. It is hoped that quantum analyses such as that presented here will be useful for treating these challenging systems. The advantage of studying proteins as models for glasses is their known structure, which makes it possible to pursue atomically detailed studies.

## REFERENCES AND NOTES

1. R. H. Austin, K. W. Beeson, L. Eisenstein, H. Frauenfelder, I. C. Gunsalus, *Biochemistry* **14**, 5355 (1975).
2. R. Elber and M. Karplus, *Science* **235**, 318 (1987).
3. R. M. Levy, O. de la Luz Rojas, R. A. Friesner, *J. Phys. Chem.* **88**, 4233 (1984).
4. J. M. Bowman, *J. Chem. Phys.* **68**, 608 (1978).
5. S. Cusack, J. Smith, J. L. Finney, B. Tidor, M. Karplus, *J. Mol. Biol.* **202**, 903 (1988).
6. B. Brooks and M. Karplus, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 6571 (1983).
7. R. Brucoleri, M. Karplus, J. A. McCammon, *Biopolymers* **25**, 1767 (1986).
8. M. Levitt, C. Sander, P. S. Stern, *J. Mol. Biol.* **181**, 423 (1985).
9. N. Go, T. Noguti, T. Nishikawa, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3696 (1983).
10. T. Nishikawa and N. Go, *Proteins* **2**, 308 (1987).
11. J. Smith, S. Cusack, B. Tidor, M. Karplus, *J. Chem. Phys.* **93**, 2974 (1990).
12. R. Elber *et al.*, *Comput. Phys. Commun.*, in press.
13. Computational details are as follows. The known crystal structure of BPTI was used to start the calculation [J. Deisenhofer and W. Steigemann, *Acta Crystallogr. B* **31**, 238 (1975)]. This structure was immersed in a large box of water molecules, and then water molecules that were strongly overlapping with the protein were removed from the system. Also removed were all water further than 3 Å from any atom in the protein. The shell of 196 water molecules that remained around the protein was then included in a molecular dynamics simulation. The system was thermalized at 300 K and maintained at that temperature for 10 ps. This structure was minimized by Powell's conjugate gradient minimizer until the norm of the gradient was  $0.003 \text{ kcal mol}^{-1} \text{ Å}^{-1}$ . This program uses the united-atom representation, so that methyl groups and methylene groups are treated as spheres. Hydrogens that participate in hydrogen bonding (for example, N-H) were included explicitly. There was no cutoff for nonbonded interactions. The force fields in this program have been widely used for ground-state structures and dynamics. The potentials in the MOL program are based on the covalent terms from the AMBER force field [S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, *J. Am. Chem. Soc.* **106**, 765 (1983)], the nonbonded interactions from the OPLS force field [W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.* **110**, 1666 (1988)], and the improper torsions from the CHARMM force field [B. R. Brooks *et al.*, *J. Comp. Chem.* **4**, 187 (1983)].
14. This is not the full quartic potential. In particular, we have neglected cubic terms with all three indices different from one another and quartic terms with an un-repeated index [ $V_{klm}^3$  and  $V_{klmnp}^4$  for example]. These averages will vanish exactly in the harmonic

- approximation, and even with anharmonicities, they are expected to be very small.
15. R. B. Gerber and M. A. Ratner, *Adv. Chem. Phys.* **70**, 97 (1988).
  16. M. A. Ratner and R. B. Gerber, *J. Phys. Chem.* **90**, 20 (1986).
  17. T. R. Horn, R. B. Gerber, M. A. Ratner, *J. Chem. Phys.* **91**, 1813 (1989).
  18. H. Romanowski, J. M. Bowman, L. Harding, *ibid.* **82**, 4155 (1985).
  19. All calculations were conducted on a Convex c240 computer. The normal modes were computed with a library parallel routine for the matrix diagonalization, a procedure that took 4 hours of computer (CPU) time. The generation of the third- and fourth-derivative matrix elements in normal coordinates was completed in 7 CPU hours. The SCF procedure was the fastest step, taking 19 s per iteration. The solutions

- to Eq. 3 were performed with a second-order finite-differences algorithm on an equispaced grid.
20. D. J. Nesbitt, *Annu. Rev. Phys. Chem.* **45**, 367 (1994); J. Laane, *ibid.*, p. 179.
21. W. Qian and S. Krimm, *Biopolymers* **34**, 1377 (1994).
22. We are grateful to the Chemistry Division of the National Science Foundation for partial support of this research. R.E. is an Allon Fellow in Israel and wishes to thank the Israel Science Foundation for support. The Fritz Haber Research Center is supported by the Minerva Gesellschaft für die Forschung, München, Germany. R.B.G. thanks the office of academic computing at the University of California at Irvine for the allocation of computational time on the Convex c240.

15 December 1994; accepted 13 March 1995

## Pressure-Tuned Fermi Resonance in Ice VII

K. Aoki,\* H. Yamawaki, M. Sakashita

Fermi resonance was observed between the OH stretch and the overtone of the OH bending modes of HDO molecules contaminated in phase VII of D<sub>2</sub>O ice over the pressure range from 17 to 30 gigapascals. An anharmonic coupling constant, which is related to the potential energy surface on which hydrogen-bonded protons oscillate, was found to range around 50 wave numbers through the resonant pressure range. Its experimentally obtained magnitude and pressure-insensitive behavior will be useful for theoretical studies of the potential energy surface and hence of the nature of hydrogen bonding in ice.

Ice is a prototypical hydrogen-bonded system, and its structural and physical properties have been a major subject of high-pressure research. More than 10 crystalline phases have been found in a pressure range up to 50 GPa and in a temperature range down to 77 K (1). Of the known phases, ice VII and VIII are particularly important for understanding the nature of the hydrogen bond. They have very closely related structures (2–4) and are stable from 2 GPa to at least 50 GPa (5). This wide range of pressure stability allows intermolecular distances to vary widely without any fundamental change in the molecular arrangement. Raman scattering (5–7), infrared absorption (8), and x-ray diffraction measurements (9, 10) have been used to investigate the high-pressure vibrational and structural behavior of these phases.

Recently, fundamental aspects of the nature of bonding in ice VII and VIII were revealed by high-pressure experiments. The pressure dependence of the hydrogen bond length in ice VIII was directly measured by neutron powder diffraction (11), and the elastic property of ice VII was precisely investigated by single crystal Brillouin spectroscopy (12). From these experimental results, the pressure dependence of the interatomic potential and the intermolec-

ular bond strength were derived. One interesting feature previously unseen in ice is resonance between the vibrational states associated with the hydrogen bonding. Such a vibrational resonance is known as Fermi resonance and is often observed in complex molecules.

Fermi resonance occurs when two vibrational modes having the same symmetry and comparable vibrational frequencies are coupled by an anharmonic term in the interatomic potential. The coupling of these two vibrational modes produces two stationary resonant states; the frequency splitting and the amplitude distribution between them are described as a function of both an anharmonic coupling constant and the frequency difference between the uncoupled initial states. Such a resonant effect was first observed in the Raman spectra of CO<sub>2</sub> in 1929 (13) and a quantum mechanical explanation was immediately given by Fermi (14). In solid CO<sub>2</sub>, the symmetric stretch and the overtone of the bending modes are coupled to form two stationary resonant states separated by 100 cm<sup>-1</sup>. The corresponding vibrational modes of H<sub>2</sub>O are far from resonance: In the infrared absorption spectra of ice VII at 2 GPa, for instance, the peaks from the stretch and the overtone of the bending modes are at about 3400 and 2900 cm<sup>-1</sup>, respectively. However, the OH stretching frequency rapidly approaches the overtone frequency with increasing pressure (8). The decrease rate,

about -20 cm<sup>-1</sup> per gigapascal, leads to an expectation that the resonance will be induced in ice VII, as it is in solid CO<sub>2</sub>, at a pressure of several tens of gigapascals. Here we describe infrared absorption measurements of ice VII up to 45 GPa at room temperature, which demonstrate the presence of pressure-tuned Fermi resonance.

Infrared spectra of HDO molecules contaminated in 99.996% D<sub>2</sub>O (Aldrich) were measured in a cylindrical diamond-anvil cell 30 mm in diameter and 20 mm thick, with an optical opening of 60° for both entrance and exit sides (15). The sample was sealed in a hole about 80 μm in diameter prepared in a 40-μm-thick metal gasket of Inconel X-750 and squeezed between opposed anvils. The pressure was determined from the shift of the fluorescence lines from ruby chips embedded in the sample (16). The pressure difference inside the sample was relatively small; at an average pressure of 36.4 GPa, for example, the measured pressures ranged from 36.3 to 36.5 GPa. Light transmitted through the unmasked sample area, a square typically 60 μm by 60 μm, was analyzed and recorded with a microscope-Fourier transform infrared spectrometer. The spectral resolution was 1.0 cm<sup>-1</sup>.

The absorption peaks of the stretching and bending vibrations in the HDO molecule were measured at pressures from 2 to 45 GPa. Over this pressure range, ice VII is known to exist stably. For ice VII, prepared just above the transition pressure of 2 GPa, a strong stretching peak was observed at about 3400 cm<sup>-1</sup> on the tail of the saturated OD stretching peaks, whereas an absorption peak associated with the overtone of the bending vibration was not observed in the expected frequency region around 3000 cm<sup>-1</sup>. The stretching peak rapidly shifted to a lower frequency with increasing pressure, reaching 3100 cm<sup>-1</sup> at 15 GPa. The overtone peak of the bending mode was then recognized as a shoulder located at about 2900 cm<sup>-1</sup>. When the pressure increased further, the two peaks approached closely enough to begin to interfere and showed anomalous changes in frequencies and peak intensities.

Figure 1 shows absorption spectra of the coupled states of the stretch and bending modes in the pressure range from 18 to 27 GPa. The intensities of the two peaks change in the opposite directions as pressure increases. The shoulder peak, initially located at 2900 cm<sup>-1</sup>, grows gradually, while the height of the large peak at 3100 cm<sup>-1</sup> decreases. Their peak heights are comparable at 22.9 GPa and then reverse at higher pressures. Finally, the original large peak is depressed to a shoulder at 26.7 GPa. These two peaks are associated with the resonant states formed by the mixing of the stretch and the overtone of the bending modes. The anom-

National Institute of Materials and Chemical Research, Tsukuba, Ibaraki 305, Japan.

\*To whom correspondence should be addressed.