Theoretical Studies of Protium/Deuterium Fractionation Factors and Cooperative Hydrogen Bonding in Peptides

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · SEPTEMBER 1995

Impact Factor: 12.11 · DOI: 10.1021/ja00143a001

3 AUTHORS:

Arthur S Edison
University of Georgia
105 PUBLICATIONS **1,910** CITATIONS
SEE PROFILE

Frank Weinhold
University of Wisconsin–Madison
199 PUBLICATIONS **28,089** CITATIONS
SEE PROFILE

John L. Markley
University of Wisconsin–Madison
576 PUBLICATIONS **18,237** CITATIONS
SEE PROFILE

VOLUME 117, NUMBER 38
SEPTEMBER 27, 1995

JOURNAL
OF THE
AMERICAN
CHEMICAL
SOCIETY

Theoretical Studies of Protium/Deuterium Fractionation Factors and Cooperative Hydrogen Bonding in Peptides

Arthur S. Edison,[†,‡] Frank Weinhold,[†,§] and John L. Markley*[,†,⊥]

Contribution from the Graduate Biophysics Program, Department of Biochemistry, and National Magnetic Resonance Facility at Madison, and Theoretical Chemistry Institute and Department of Chemistry, University of Wisconsin at Madison, Madison, Wisconsin 53706

Received January 13, 1995[⊗]

Abstract: We have calculated fractionation factors (ϕ) for a water cluster and a wide range of peptide clusters. We find that low fractionation factors can occur when *both* charged interactions *and* cooperative hydrogen-bonded networks are present. Correlations between fractionation factors and hydrogen-bonded N−O distances or N−H bond lengths have been found which are nonlinear and show maximum values of fractionation factors at N−O distances of about 2.80 Å. These calculations support the wide range (0.28−1.47) of fractionation factors found in proteins [Loh, S. N.; Markley, J. L. *Biochemistry* **1994**, 33, 1029−1036].

I. Introduction

Recently, Loh and Markley have reported experimental NMR measurements of protium/deuterium fractionation factors, ϕ, of most of the exchangable amide hydrogens in the protein staphylococcal nuclease.[1,2] The fractionation factor is defined as the equilibrium constant for the reaction

$$PH + SD \rightleftharpoons PD + SH \qquad (1)$$

where, for proteins, P is an amide nitrogen of the peptide and the S refers to a bulk $H_2O/D_2O/DHO$ solvent. The unusual feature found by Loh and Markley was that, across the amino acid sequence, ϕ spans the large range of values from 0.28 to 1.47,[1] which are representative of both the lowest and highest fractionation factors found for all types of molecules (for example, see refs 3−13). Particularly curious are the extremely low values of ϕ. Loh and Markley found 24 cases with $\phi \leq 0.70$ from both the ligated (complex with thymidine-3′,5′-bisphosphate and calcium) and unligated forms of staphylococcal nuclease.[1] The only types of molecules known to exhibit extremely low fractionation factors are those involved in very strong, symmetric hydrogen bonds which usually involve charged groups.[6,11,14] For example, Graul and co-workers reported gas phase experimental measurements of the fraction-

[†] Graduate Biophysics Program.
[‡] Current address: Department of Zoology, University of Wisconsin, Madison, WI 53706.
[§] Theoretical Chemistry Institute and Department of Chemistry.
[⊥] National Magnetic Resonance Facility.
[⊗] Abstract published in *Advance ACS Abstracts,* September 1, 1995.

(1) Loh, S. N.; Markley, J. L. In *Techniques in Protein Chemistry*; Angeletti, R., Ed.; Academic Press, 1993; Vol. 4, pp 517−524.
(2) Loh, S. N.; Markley, J. L. In *Techniques in Protein Chemistry*; Angeletti, R., Ed.; Academic Press, 1993; Vol. 4, pp 517−524.
(3) Kresge, A. J.; Allred, A. L. *J. Am. Chem. Soc.* **1963**, *86*, 1541−1541.
(4) Heinzinger, K.; Weston, R. E., Jr. *J. Phys. Chem.* **1964**, *68*, 744−751.
(5) Cleland, W. W. *Methods Enzymol.* **1980**, *64*, 105−125.
(6) Kreevoy, M. M.; Liang, T. M. *J. Am. Chem. Soc.* **1980**, *102*, 3315−3322.
(7) Kurz, J. L.; Myers, M. T.; Ratcliff, K. M. *J. Am. Chem. Soc.* **1984**, *106*, 5631−5634.
(8) Jarret, R. M.; Saunders, M. *J. Am. Chem. Soc.* **1985**, *107*, 2648−2654.
(9) Jarret, R. M.; Saunders, M. *J. Am. Chem. Soc.* **1986**, *108*, 7549−7553.
(10) Larson, J. W.; McMahon, T. B. *J. Am. Chem. Soc.* **1986**, *108*, 1719−1720.
(11) Graul, S. T.; Brickhouse, M. D.; Squires, R. R. *J. Am. Chem. Soc.* **1990**, *112*, 631−639.
(12) Cleland, W. W.; Kreevoy, M. M. *Science* **1994**, *264*, 1887−1890.

0002-7863/95/1517-9619$09.00/0

ation factors for the reaction

$$[H,D]_2O-H-[H,D]DO^+ \rightleftharpoons [H,D]_2O-D-[H,D]HO^+ \quad (2)$$

that range from 0.66 to 0.80, increasing with the total number of deuterons.[11] The notation [H,D] refers to a site with either an H or D atom. These results are consistent with other experimental values for protonated water solutions where ϕ ranges from 0.69 to 0.79 under different experimental conditions,[3,4,7,10] We recently completed theoretical calculations of ϕ for eq 2 using *ab initio* geometries and vibrational frequencies and found a range of values between 0.46 (at the single split-valence double-ζ RHF/3-21G level of theory) and 0.64 (at the highly extended RHF/6-311++G* level of theory).[15] Del Bene has calculated hydrogen bond enthalpies for the $O_2H_5^+$ cluster at several basis sets and levels of theory and obtained a value of −32.4 kcal/mol at the highest level of theory she examined.[16] This number agrees quite well with an experimental measurement of −31.8 kcal/mol for the hydrogen bond strength of $O_2H_5^+$.[17] Thus, it is clear from many experimental and theoretical results that, at least for $O_2H_5^+$ clusters, the low values of ϕ (not as low as some of the protein values) correspond to very strong hydrogen bonds.

Following Kreevoy,[18] fractionation factors can be understood qualitatively by considering effective one-dimensional potential energy curves of differing curvature. Figure 1 illustrates three different cases corresponding to eq 1: peptide site X has a stronger overall force constant for X−[H,D] than the solvent S has for S−[H,D], so, to minimize the total energy, the equilibrium constant (fractionation factor) for eq 1 will be greater than unity; peptide site Y has a force constant equal to that of the solvent, so the fractionation factor is unity; and peptide site Z has a force constant less than that of the solvent and consequently has a fractionation factor less than unity. These potential curves can be misleading, because they are not meant to correspond to any one vibrational mode,[19] but they provide a qualitative visual interpretation of the physical origin of fractionation factors.

Our goal in this paper is twofold. First, we want to explain theoretically the wide range of amide fractionation factors observed in proteins. Second, we want to find a physical basis (e.g. structural correlation) for the very low fractionation factors in proteins.[1] To accomplish these goals, we have performed *ab initio* calculations to obtain optimized geometries and vibrational frequencies for a large number of amides, amide clusters, and water clusters. Early in our investigations we found that relatively small amides and clusters would be insufficient to explain the complete range of fractionation factors from proteins, so we were forced to examine some rather large molecular clusters. As a result of the large number and size of molecular clusters, we have limited our studies to the RHF/3-21G level of theory (with a few calculations also repeated at the RHF/6-31+G* level as a check). Clearly, RHF/3-21G will only provide qualitative results, but other studies have shown that this level of theory can provide surprisingly good results in calculations of fractionation factors.[14,15,20−22] More importantly, we are attempting to calculate properties of proteins, so the approximations of using simple fragments as models for various regions of a protein will probably outweigh other theoretical approximations. The paper is organized as follows. In section II, we describe the calculation of fractionation factors and how we introduce solvent into eq 1. In section III we present the results of the *ab initio* calculations and the calculated fractionation factors, and in section IV, we discuss the major results. We close by suggesting further theoretical and experimental studies that might help unravel some of the complexities of the experimental protein results.

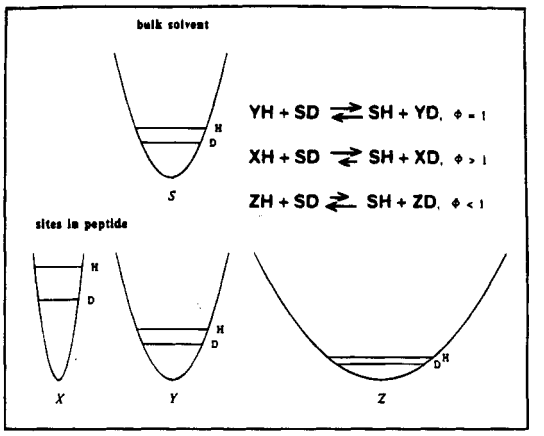

Figure 1. Schematic representations of abstract, one-dimensional potential energy wells for the solvent (S) and three peptide amides (X, Y, and Z) with different force constants relative to the solvent. The zero-point energies for H and D are shown for all cases. In the first case, X has a stronger force constant than S, with the result that deuterium will preferentially migrate onto X. The third case shows the reverse situation, in which Z will become enriched in protium relative to S. In the second case, H and D will be partitioned equally between the solvent and peptide site.

II. Methods

Ab initio calculations at the RHF/3-21G and RHF/6-31+G* levels were used to obtain optimized geometries and analytical frequencies for the molecules and clusters shown in Figure 2.[23] The formamide clusters (FA_n) were constrained to a plane, and have some small imaginary frequencies of about 20−30 cm^{-1} (one for FA_2, two for FA_3, three for FA_4, and two for FA_6). We also fully optimized FA_2 by removing the planar constraint which led to the negative frequencies and obtained virtually the same result for fractionation factors as before, probably because frequencies lower than about 50 cm^{-1} contribute very little to ϕ (see eq 3 below). All the other structures were fully optimized and yielded no imaginary frequencies. The optimized Cartesian

(13) Frey, P. A.; Whitt, S. A.; Tobin, J. B. *Science* **1994**, *264*, 1927−1930.

(14) Weil, D. A.; Dixon, D. A. *J. Am. Chem. Soc.* **1985**, *107*, 6859−6865.

(15) Edison, A. S.; Markley, J. L.; Weinhold, F. *J. Phys. Chem.* In press. We also used MP2 to estimate electron correlation effects, but found values that were much too low (0.30−0.49, depending on basis set), most likely because the MP2 hydrogen bond lengths were too short.

(16) Del Bene, J. E. *J. Phys. Chem.* **1987**, *8*, 810−815.

(17) Meot-Ner (Mautner), M. *J. Am. Chem. Soc.* **1986**, *108*, 6189−6197.

(18) Kreevoy, M. M. In *Isotopes in Organic Chemistry*; Buncel, E., Lee, C. C., Eds.; Elsevier Scientific Publishing Co.: New York, 1976; Vol. 2, pp 1−31.

(19) The vibrational potentials shown in Figure 1 are usually assumed to correspond to the A−[H,D] or B−[H,D] stretching frequencies, but, as we show in the Discussion section, other vibrational modes can contribute substantially to an overall difference in total zero point energies. Thus, the schematic potentials in Figure 1 should probably be thought of as effective "total" modes corresponding to the total zero point energy in the molecule.

(20) Hout, R. J., Jr.; Wolfsberg, M.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 3296−3298.

(21) Saunders, M.; Laidig, K. E.; Wolfsberg, M. *J. Am. Chem. Soc.* **1989**, *111*, 8989−8994.

(22) Williams, I. H. *J. Phys. Org. Chem.* **1990**, *3*, 181−190.

(23) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley & Sons, Inc.: New York, 1986.

(24) Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Degrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. *Gaussian 92*, 1992, Revision A, Gaussian Inc., Pittsburgh, PA.

coordinates and force constant matrix from the Gaussian 92[24] output files were used as input for the computer program QUIVER[21] which calculates the ratios of the reduced isotopic partition functions for PD/PH and SD/SH from eq 1.[25] The reduced isotopic partition function, $(s/s')f$, is given by

$$\frac{s}{s'}f = \prod_{i}^{3N-6} \frac{u_i}{u'_i} \frac{e^{-u_i/2}/(1-e^{-u_i})}{e^{-u'_i/2}/(1-e^{-u'_i})} \quad (3)$$

where $u_i = h\nu_i/kT$, ν_i are the $3N-6$ vibrational frequencies, s is the symmetry number, and the prime refers to the lighter isotope.[25] All of the symmetry numbers in this paper are unity and thus will be neglected here.[26] The equilibrium constant (fractionation factor) for eq 1 is given by

$$\phi = \frac{f(\text{PD/PH})}{f(\text{SD/SH})} \quad (4)$$

For f(PD/PH), we simply replace a proton with a deuteron at a given site of one of the peptides. Finding f(SD/SH) for water is not straightforward, because the molecular properties of bulk water are not well understood. With extensive hydrogen bonding, a water monomer is not likely to make a significant contribution in bulk liquid. Preliminary calculations of the thermodynamic properties of water using water cluster sizes from 1 to 18 have indicated that large cyclic water clusters (such as the pentamer, hexamer, octamer, ...) are dominant (>90%) species in the liquid state (ref 27 and Weinhold, unpublished results). Theoretical predictions of this quantum cluster equilibrium (QCE) model are in excellent agreement with experimentally measured quadrupole coupling constants, asymmetry parameters, and their associated temperature dependence for water[27] as well as neat liquid formamide.[28,29] The larger cyclic n-mers ($n \geq 5$) have rather similar electronic environments for each monomer, with strainless (near-linear) H-bonds and comparable cooperative enhancements (section IV), and they are found to play essentially equivalent thermodynamic roles in the QCE description. Therefore, we have chosen to use the cyclic water hexamer as the representative molecule for S in eq 1. The water hexamer has two distinct hydrogen sites, hydrogen bonded and free, with values of f(SD/SH) equal to 17.85 and 14.85, respectively. In calculations of fractionation factors from eq 4, we use the average of these two sites, giving f(SD/SH) = 16.35. Substantially similar values would be obtained for any of the other large cyclic clusters dominating the liquid phase.

III. Results

In this study, we have calculated several examples from three categories of peptide models: (1) formamide (FA_n) clusters (n = 2, 3, 4, and 6), (2) δ-valerolactam (DVL) clusters (monomer, dimer, monomer + 1 water, and monomer + 2 waters), and (3) monomers and dimer of two amino acid derivatives, *cis-N*-(2-carboxyethyl)formamide (AA-I) and 3-ammoniumpropanamide (AA-II). The optimized structures and abbreviations are shown in Figure 2. Each formamide monomer in the various clusters is labeled from "a" to "d" from the N-terminal to C-terminal end of the clusters. FA_6 is a cyclic structure with C_3 symmetry (Figure 2), so it has two unique monomers, "a" and "b". Table 1 gives the hydrogen bonded N−O and N−H bond lengths and N−H−O angles, and Table 2 gives the calculated fractionation factors and all the N−H bond lengths (hydrogen bonded and free) for each N−L (L = H or D) site in each of the molecules and clusters shown in Figure 2.

Table 1. Table of RHF/3-21G Hydrogen-Bonding Parameters[a]

molecule	R_{NO} (Å)	r_{NH} (Å)	θ (deg)[b]
FA_2	2.920	1.003	
FA_3 (a−b)	2.865	1.006	
FA_3 (b−c)	2.880	1.007	
FA_4 (a−b)	2.851	1.007	
FA_4 (b−c)	2.827	1.010	
FA_4 (c−d)	2.871	1.008	
FA_6 (a−b)	2.800	1.015	
FA_6 (b−a)	2.743	1.003	
DVL + H_2O	2.795	1.011	145.8
DVL + $2H_2O$	2.773	1.023	178.1
2DVL	2.830	1.017	176.7
AA-I	2.572	1.028	143.0
AA-II	2.542	1.058	143.9
AA-I + AA-II			
intramolecular AA-I	2.447	1.099	151.3
intramolecular AA-II	2.461	1.128	152.4
intermolecular	2.604	1.000	157.5

[a] The molecule abbreviations are explained in the text and in Figure 2. [b] θ is the angle N−H−O.

IV. Discussion

Cooperativity. The most striking result of the fractionation factor calculations is the large change, upon dimerization, in the intramolecular hydrogen bonds of AA-I and AA-II (Figure 2). Upon forming a dimer, ϕ drops from 1.20 to 0.64 and from 1.10 to 0.59 for AA-I and AA-II, respectively (Table 2). The intramolecular values in the dimer (0.59 and 0.64) are the lowest fractionation factors that we have calculated in a peptide. All other fractionation factors in AA-II remain relatively constant between the monomer and dimer. This change in fractionation factor is also reflected in the intramolecular N−O bond distances which drop, upon dimerization, from 2.572 to 2.447 Å and from 2.542 to 2.461 Å for AA-I and AA-II, respectively (Table 1). Likewise, the N−H bond lengths increase from 1.028 to 1.099 Å and from 1.058 to 1.128 Å for AA-I and AA-II. These changes are the result of a cooperative enhancement of hydrogen bond strengths, mediated through the peptide bond, which can be represented in terms of resonance between the structures shown in Figure 3. Without the intermolecular hydrogen bond, no such cooperative enhencement is possible.

A second cooperative effect is seen in the calculated shortening of the intermolecular N−O bonds and the lengthening of the hydrogen-bonded N−H bonds in formamide clusters as the number of monomers is increased. FA_2 has N−O and (H-bonded) N−H bond lengths of 2.920 and 1.003 Å, FA_4 has a central N−O and N−H bond lengths of 2.827 and 1.010 Å, and the three shortest H bonds in FA_6 have N−O and N−H distances of 2.800 and 1.015 Å, respectively. However, the calculated fractionation factors are remarkbly insensitive to these changes, as will be discussed below.

Hydrogen bonding can be understood in terms of natural bond orbitals (NBOs) by considering electron donor → acceptor relationships.[30,31] In terms of NBOs, a hydrogen bond *donor* has an electron lone pair (n) to donate to the antibond (σ^*) of the hydrogen bond *acceptor*, so an amide N−H hydrogen bonded to water or a carbonyl oxygen is an $n_O \rightarrow \sigma^*_{NH}$ interaction. A donor → acceptor relationship is completely analogous to the qualitative "electron arrow pushing" so important to organic chemistry and related to Sidgwick's

(23) Bigeleisen, J.; Mayer, M. G. *J. Chem. Phys.* **1947**, *15*, 261−267.
(24) The symmetry of the reference water hexamer is actually 3, but since this is only meant to approximate the complicated mixture of water species in solution, there is no physical justification for expecting a symmetry number effect from the solvent.
(25) Ludwig, R.; Weinhold, F.; Farrar, T. C. *J. Chem. Phys.* In press.
(26) Ludwig, R.; Weinhold, F.; Farrar, T. C. *J. Chem. Phys.* **1995**, *102*, 5118−5125.
(27) Ludwig, R.; Weinhold, F.; Farrar, T. C. *J. Chem. Phys.* In press.
(28) Foster, J. P.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, *102*, 7211−7218.
(29) Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899−926.

Table 2. Calculated Fractionation Factors[a] for the Reaction

$$P{-}H + HDO(H_2O)_5 \rightleftharpoons P{-}D + (H_2O)_6$$

molecule[b] (site)	ϕ	r_{NH} (Å)
FA_2 (*t* a)	1.15	0.996
FA_2 (*c* a)	1.15	0.999
FA_2 (*t* b)*	1.33	1.003
FA_2 (*c* b)	1.15	0.998
FA_3 (*t* a)	1.16	0.996
FA_3 (*c* a)	1.16	0.999
FA_3 (*t* b)*	1.34	1.006
FA_3 (*c* b)	1.17	0.999
FA_3 (*t* c)*	1.34	1.007
FA_3 (*c* c)	1.15	0.998
FA_4 (*t* a)	1.16	0.996
FA_4 (*c* a)	1.16	0.999
FA_4 (*t* b)*	1.34	1.007
FA_4 (*c* b)	1.17	0.999
FA_4 (*t* c)*	1.34	1.010
FA_4 (*c* c)	1.17	0.999
FA_4 (*t* d)*	1.34	1.008
FA_4 (*c* d)	1.15	0.998
FA_6 (*t* a)*	1.33	1.015
FA_6 (*c* a)	1.17	0.998
FA_6 (*t* b)*	1.32	1.003
FA_6 (*c* b)	1.16	0.997
DVL	1.26	1.000
DVL + H_2O*	1.49	1.011
DVL + $2H_2O$*	1.47	1.023
2 DVL*	1.46	1.017
AA-I*	1.20	1.028
AA-II [NH_2 H(*E*), *t*]	1.20	0.998
AA-II [NH_2 H(*Z*), *c*]	1.21	1.002
AA-II [NH_3^+]*	1.10	1.058
AA-II [NH_3^+]	1.36	1.015
AA-II [NH_3^+]	1.35	1.015
AA-I + AA-II (intra AA-I)*	0.64	1.099
AA-I + AA-II (AA-II [NH_2 H(*E*), *t*])*	1.18	1.037
AA-I + AA-II (AA-II [NH_2 H(*Z*), *c*])	1.22	1.000
AA-I + AA-II (AA-II [NH_3^+])*	0.59	1.128
AA-I + AA-II (AA-II [NH_3^+])	1.34	1.011
AA-I + AA-II (AA-II [NH_3^+])	1.34	1.012

[a] The values were obtained using force constants and geometries from RHF/3-21G calculations. The numbers were calculated on the basis of an average value of 16.35 for the half of the reaction involving the water hexamer. *Cis* (*c*) and *trans* (*t*) are with respect to the oxygen. See Figure 2 for labeling. [b] An asterisk indicates an N−[H,D] involved in a hydrogen bond.

resonance-theoretic description of solvation.[32] Cooperative effects in hydrogen bonding[31,33−36] can be understood with the help of Figure 4 where we show the n_O and σ^*_{NH} NBOs for the intramolecular hydrogen bond of AA-I in the monomer (A) and the dimer (B). The orbital overlap (which is proportional to the strength of the $n_O \rightarrow \sigma^*_{NH}$ interaction) increases dramatically upon dimer formation. The increase in the $n_O \rightarrow \sigma^*_{NH}$ orbital overlap upon dimer formation is a result of a similar $n_O \rightarrow \sigma^*_{NH}$ interaction from *inter*molecular hydrogen bond formation between the lone pair of the carbonyl oxygen of AA-I and the NH antibond of AA-II. By forming the intermolecular hydrogen bond, the lone pair of the AA-I oxygen is delocalized toward the AA-II NH, and the NH antibond is polarized toward the oxygen. This polarization results in a stronger $n_N \rightarrow \sigma^*_{CO}$ interaction across the peptide bond of AA-I which, in turn, makes the σ^*_{NH} a stronger electron acceptor. As a result, the intramolecular hydrogen bond ($n_O \rightarrow \sigma^*_{NH}$) depicted in Figure 4B is stronger than that in Figure 4A. From second-order perturbation theory estimates,[31] the intramolecular $n_O \rightarrow \sigma^*_{NH}$ interaction energies are 45.64 and 110.27 kcal/mol for the AA-I monomer and AA-I + AA-II dimer, respectively.

Changes in Fractionation Factors upon Hydrogen Bonding. Every N−[H,D] site, except the charged groups in AA-I and AA-II discussed above, has a higher fractionation factor (becomes enriched in deuterium relative to protium) when the N−[H,D] is hydrogen bonded than when it is free. This result is counter-intuitive, because the N−[H,D] stretching frequency drops with hydrogen bonding, and according to Figure 1,[18] protium would preferentially migrate to the bond with a lower frequency. The answer lies in the modes other than the N−[H,D] stretching frequency. When a hydrogen bond is formed, the stretching frequency is reduced, but the in- and out-of-plane bending modes are *stiffened*. As a result, the fractionation factor will represent a balance between decreasing the stretching frequency and increasing the bending frequencies of the hydrogen bonded group (Kreevoy, personal communication). The N−H group in peptides contributes to many normal modes, so that a direct comparison of normal modes with and without hydrogen bonds is difficult. Although the fractionation factors given in Table 2 represent contributions from every vibrational mode, some qualitative comparisons can be useful. For example, the N−H stretching frequency for the DVL monomer is 3785 cm^{-1}; the corresponding frequency for the DVL + H_2O cluster is 3647 cm^{-1}. However, upon H-bond formation, the mode most closely associated with the N−H out-of-plane bend increases frequency from 776 (DVL) to 1086 cm^{-1} (DVL + H_2O).

Correlations of Fractionation Factors with N−O and N−H Distances. For fractionation factors to have practical value in studies of protein structures, there must be a geometrical correlation. Figure 5 shows calculated fractionation factors vs N−O distance for hydrogen-bonded groups (A) and N−H distance for all groups (B). The interesting feature of parts A and B of Figure 5 is that the fractionation factor appears to go through a maximum of about 1.4 at N−O and N−H distances of 2.80 and 1.01 Å, respectively. The effect appears to be greater for the N−H distances (Figure 5B), because the non-hydrogen-bonded N−H groups (that have no corresponding hydrogen-bonded N−O distance) are included; these show lower fractionation factors than weakly hydrogen-bonded groups. The N−H distance, which normally is not resolved in protein crystal structures, might provide the strongest correlation with fractionation factors. In light of the previous discussion about the effects of hydrogen bonding, this maximum should come as no surprise. Recall that, although the formamide clusters showed large geometrical changes as cluster sizes increased, the fractionation factors were relatively constant (Table 2); Figure 5 shows that the maximum value of ϕ corresponds to the general region of the N−O and N−H distances in the formamide clusters. Therefore, even though fairly large geometrical changes are observed as a result of cooperativity, the apparent functions for the fractionation factor are reasonably flat and insensitive in these regions.

V. Conclusions and Future Studies

We have calculated fractionation factors for a number of different types of peptides with and without hydrogen bonds, and of all the complexes studied, only one shows low

(30) Sidgwick, N. V. *The Electronic Theory of Valency;* Oxford University Press: London, 1929.
(31) Guo, H.; Karplus, M. *J. Phys. Chem.* **1994**, *98*, 7104−7105.
(32) Suhai, S. *Int. J. Quantum Chem.* **1994**, *52*, 395−412.
(33) King, B. F.; Weinhold, F. *J. Chem. Phys.* **1995**, *103*, 333−347.
(34) King, B. F.; Farrar, T. C.; Weinhold, F. *J. Chem. Phys.* **1995**, *103*, 348−352.

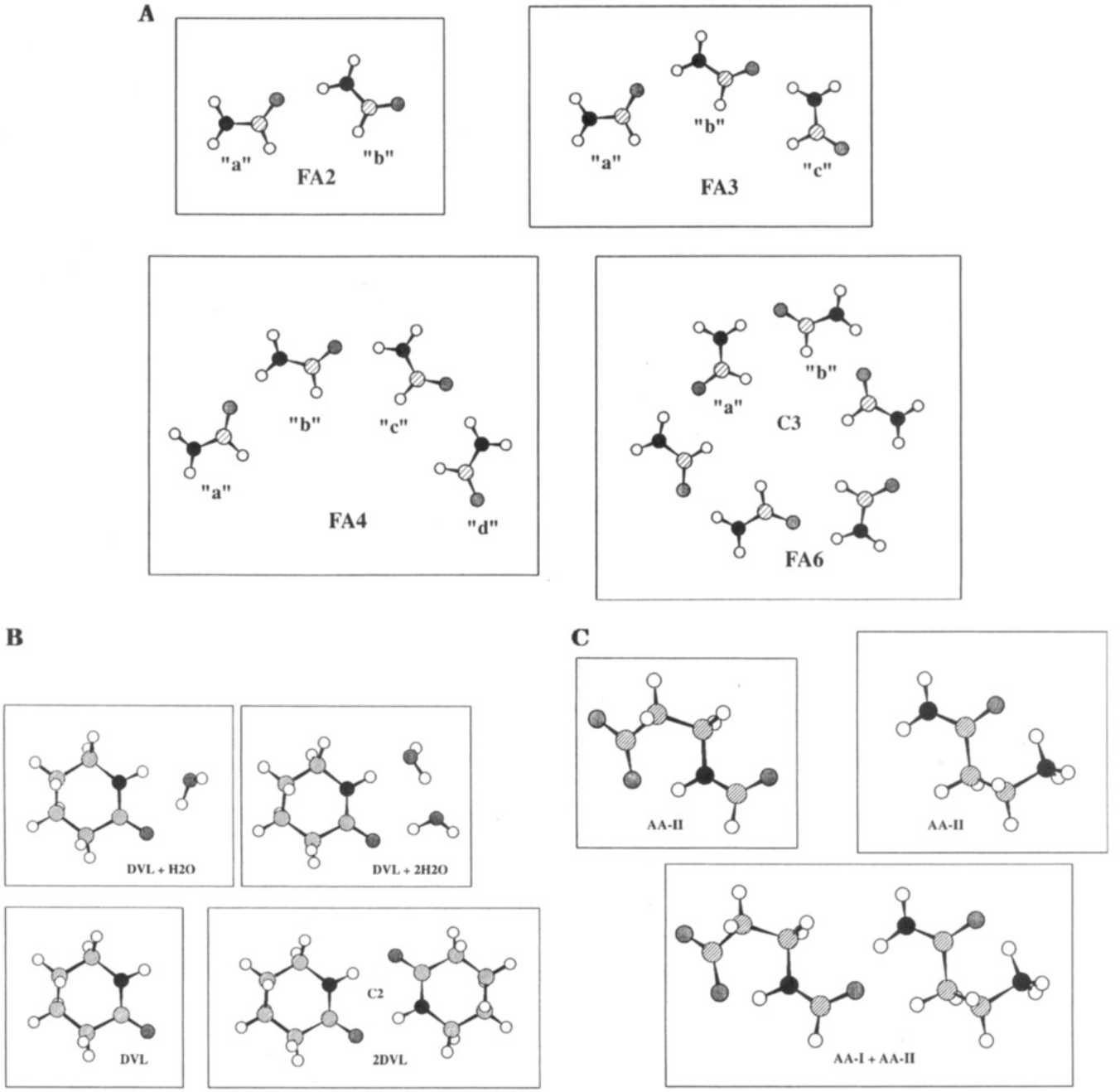

Figure 2. Optimized RHF/3-21G geometries and nomenclature for the molecules and clusters used in this study. Hydrogen atoms are unshaded, carbon atoms are striped, nitrogen atoms are black, and oxygen atoms are gray. The formamide clusters (FA_n) were constrained to planarity and showed some (one for FA_2, two for FA_3, three for FA_4, and two for FA_6) small negative frequencies (about 20 cm^{-1}) which correspond to low-frequency out-of-plane oscillations of the clusters. All other geometries shown had all real frequencies.

fractionation factors (AA-I + AA-II, Figure 2). Although the number of examples is currently insufficient to make sweeping statements, our calculations suggest that low fractionation factors seen in proteins[1] likely result from some type of charged interactions, possibly enhanced by cooperative networks of hydrogen bonds. Hydrogen-bonding cooperativity might contribute significantly to protein structures and folding. The α-helix and β-sheet are both characterized by extended networks of hydrogen bonds between peptide groups, and the strengths of the hydrogen bonds should increase (toward an asymptotic limit) with the extent of the hydrogen-bonded network. It is of interest that out of 15 amides hydrogen bonded to aspartic or glutamic acids in the ligated and unligated forms of staphyococcal nuclease, 5 have $\phi < 0.7$, 10 have $\phi < 0.83$, and all are less than 1.05. In addition, the lowest fractionation factor (0.28 in T120) is for hydrogen bonded to aspartic acid.[1]

Our results also suggest that any correlation between hydrogen-bonded distances and ϕ should be nonlinear and exhibit a maximum. Currently, no clear correlation between N−O distances and ϕ is seen in the experimental data from staphylococcal nuclease,[1] but factors such as hydrogen bond angle "strain" and dielectric changes across the amino acid sequence are also likely to be important and may obscure any simple correlation with distance.

Future studies should be aimed at elucidating the influence of other electronic and geometrical variables. For example, the dependence on charge could be investigated by studying various ion−neutral complexes as well as higher polyionic species. In addition, various ternary complexes could be studied that correspond to breaking the covalent linkages holding the intramolecular hydrogen bonds in AA-I and AA-II, thus allowing optimal hydrogen-bonding arrangements. Either effect

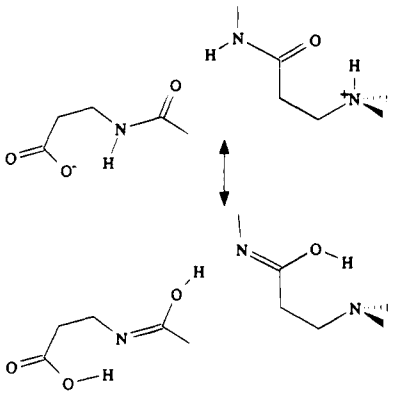

Figure 3. Resonance structures of the AA-I + AA-II complex that serve to illustrate the cooperative nature of hydrogen bonding.

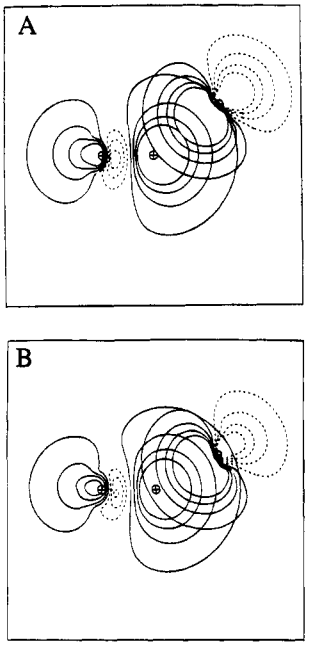

Figure 4. Orbital contour plots of the AA-I intramolecular hydrogen-bonding n_O (lone pair on oxygen) and σ^*_{NH} (antibond on N−H) NBOs of AA-I monomer (A) and AA-I + AA-II (B). The large increase in overlap in B results in a shorter, stronger hydrogen bond and leads to a large reduction of the fractionation factor for the N−[H,D]. This change from A to B represents a cooperative enhancement of the hydrogen bond due to the formation of an intermolecular hydrogen bond as described in the text.

may lead to significantly lower ϕ values than those (0.59 and 0.64) found in the current work for the AA-I and AA-II complex.

Our theoretical model also contains implications that could be subjected to direct experimental tests. For example, one could examine the effect of placing an aspartic or glutamic acid at the N-terminal end of an α-helix. Our results suggest that the N−[H,D] group hydrogen bonded to a carboxylate will have a much lower fractionation factor than the same group without

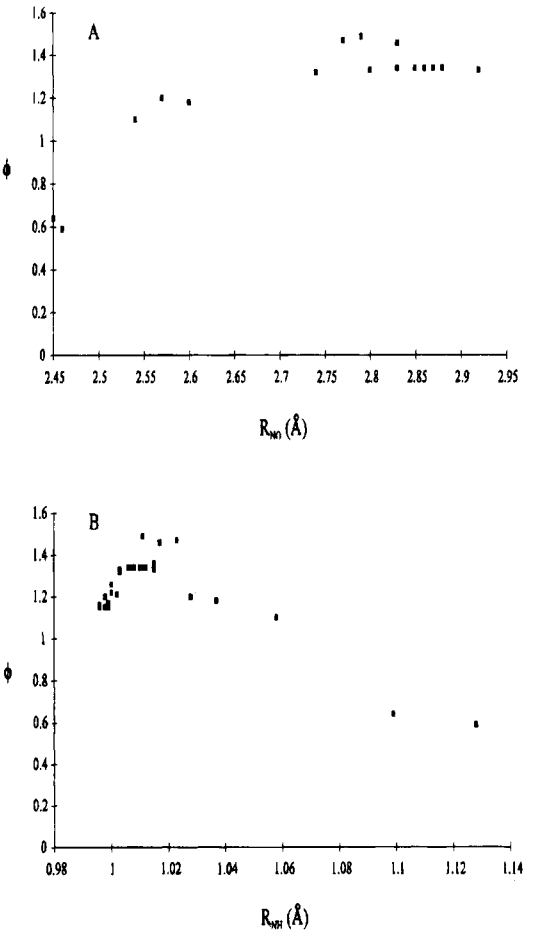

Figure 5. Fractionation factors (ϕ) vs (A) hydrogen-bonded N−O and (B) N−H distances. The apparent maxima can be understood by the initial increase of ϕ upon weak hydrogen bonding and subsequent decrease as the hydrogen bonding becomes stronger.

the acid. The measured fractionation factor in such a system should also be highly dependent upon pH, with higher values of ϕ at low pH where the carboxyl group is neutral and lower values at moderate to high pH where it is negatively charged. The results here suggest that when very small ϕ values are found in a protein, the N−H bond distance should be correspondingly long (Figure 5B). It may be possible to verify this by neutron diffraction analysis.

Acknowledgment. We thank Dr. Stewart Loh for experimental results and advice and Profs. W. Wallace Cleland and Maurice Krevoy for helpful discussions. Prof. Martin Saunders kindly provided the source code for QUIVER which was used in calculations of the reduced isotopic partition functions. *Ab initio* calculations were performed with Gaussian 92 installed on a Kubota/Stardent TITAN in the Chemistry Department and a Cray Y-MP at Cray Research, Inc. A.S.E. was a trainee of an NIH Molecular Biophysics Training Grant (GM08293). J.L.M. has support from NIH Grant No. RR02301.

JA950123X