Van der Waals Interactions Involving Proteins

Charles M. Roth, Brian L. Neal, and Abraham M. Lenhoff
Center for Molecular and Engineering Thermodynamics, Department of Chemical Engineering, University of Delaware, Newark, Delaware 19716 USA

ABSTRACT Van der Waals (dispersion) forces contribute to interactions of proteins with other molecules or with surfaces, but because of the structural complexity of protein molecules, the magnitude of these effects is usually estimated based on idealized models of the molecular geometry, e.g., spheres or spheroids. The calculations reported here seek to account for both the geometric irregularity of protein molecules and the material properties of the interacting media. Whereas the latter are found to fall in the generally accepted range, the molecular shape is shown to cause the magnitudes of the interactions to differ significantly from those calculated using idealized models, with important consequences. First, the roughness of the molecular surface leads to much lower average interaction energies for both protein-protein and protein-surface cases relative to calculations in which the protein molecule is approximated as a sphere. These results indicate that a form of steric stabilization may be an important effect in protein solutions. Underlying this behavior is appreciable orientational dependence, one reflection of which is that molecules of complementary shape are found to exhibit very strong attractive dispersion interactions. Although this has been widely discussed previously in the context of molecular recognition processes, the broader implications of these phenomena may also be important at larger molecular separations, e.g., in the dynamics of aggregation, precipitation, and crystal growth.

INTRODUCTION

Protein molecules interact with like proteins, with different proteins, and with other materials by a number of mechanisms, including electrostatics, van der Waals (dispersion) forces, and solvation forces. Much of the current understanding of such interactions has been inferred from observations of bulk thermodynamic behavior, such as solubility, osmotic pressure, and adsorption (Melander and Horvath, 1977; Vilkner et al., 1981; Norde, 1986; Haynes et al., 1992). More recently, however, intermolecular forces have been measured directly for many materials, including proteins, by techniques such as the surface forces apparatus (Afshar-Rad et al., 1987; Leckband et al., 1994), atomic force microscopy (Ducker et al., 1991), and osmotic stress (Prouty et al., 1985; Colombo et al., 1992). Interpretation of interaction data is usually made by assuming additive contributions from the different types of forces believed to be important, but some of these forces are understood incompletely at the molecular level, particularly for complex macromolecules such as proteins. As more measurements of this type emerge, it is important to examine the manifestation of the different forces in proteins, to ensure both meaningful and accurate interpretation of experimental data and realistic prediction of macroscopic behavior of proteins via incorporation of the interactions in solution thermodynamic models. Because experimental decomposition of the protein interactions into their individual contributions is not feasible, a solid foundation is needed for predicting them theoretically; such predictions should be based, as far as possible, on explicit protein structural information and solution conditions. Structures obtained for many proteins by means of x-ray crystallography and NMR have been utilized in various studies seeking to understand intramolecular, and to some extent intermolecular, interactions involving proteins, within the framework of models that incorporate explicitly each atom in the protein (Lu and Park, 1990; Northrup et al., 1990; Lim and Herron, 1992). Unfortunately, the number of atoms involved makes this approach very expensive computationally for studying intermolecular processes involving proteins, especially if the solvent is treated atomistically as well. A common alternative is to model a protein in a colloidal context, taking the simplified view of a protein molecule as a charged, hard sphere or ellipsoid (Vilkner et al., 1981; Jeon and Andrade, 1991; Haynes et al., 1992; Roth and Lenhoff, 1993, 1995), an approach that obviously neglects most geometric information apart from the molecular size. An intermediate route is to retain the colloidal approach but to include some structural characteristics of the molecules. A considerable amount of work has been done in this regard through the use of cavity dielectric models (Kirkwood, 1934) in molecular electrostatic computations, which have been used to describe molecular events including electron transport, enzyme-substrate steering, shifts in titration behavior, and adsorption to surfaces (Warwicker and Watson, 1982; Klapper et al., 1986; Bashford and Karplus, 1991; Yoon and Lenhoff, 1992; Roush et al., 1994). In these applications, continuum electrostatic equations are employed, with structural considerations incorporated via the description of the protein boundary and...
its internal charge distribution derived from structural data; the crystallographic structure is assumed to be an adequate approximation to the structure in solution.

Here we characterize van der Waals (dispersion) interactions involving proteins in a similar fashion. Dispersion forces are typically of short range relative to, for example, electrostatic forces. Their importance is a consequence of the fact that, for a protein molecule interacting with another protein molecule or with a surface in water, they are attractive under most conditions of interest: their strength increases sharply as the intervening distance decreases. As a result, accounting for dispersion interactions must be included in the interpretation of force measurements at short range and in analyses of aggregation in solution or of adsorption to a surface.

The van der Waals interactions of atoms comprise a sum of Keesom, Debye, and London contributions, the energy of each of which decays as the inverse sixth power of separation distance for nonretarded (i.e., short-range) interactions (Prausnitz et al., 1986). For dipole moments less than about 1 Debye, the largest contribution is made by London dispersion forces, for which the prefactor is proportional to the product of the atomic polarizabilities. The van der Waals interaction between two macroscopic bodies is the resultant of the interactions of their constituent atoms.

Most analyses of van der Waals energetics in colloidal bodies utilize the simple approach of Hamaker (1937), who postulated that the contributions of the individual atoms to the overall dispersion energy should be approximated by pairwise summation. For the case of two macroscopic bodies comprising materials that are isotropic with respect to their polarizabilities, this corresponds to an integration over their volumes, and their interaction energy across a vacuum is thus found in the Hamaker analysis by integrating over the volumes $V_1$ and $V_2$ of the two bodies:

$$\Delta F_{12} = -\frac{A_{12}}{2} \int \frac{1}{r_{12}^6} \, dv_1 \, dv_2.$$  

(1)

Here $r_{12}$ is the distance between volume elements $v_1$ and $v_2$ in the two bodies, $A_{12}$ is the Hamaker constant, dependent only on the polarizability and number density of atoms in the two media, and the change in free energy $\Delta F_{12}$ is relative to that of the bodies at infinite separation. For two bodies interacting across a third medium, such as water, the same result applies, but the value of the Hamaker constant, now denoted as $A_{132}$, is modified to account for the polarizability of the intervening medium.

The Hamaker approach correctly associates the material property $A_{12}$ with the mutual polarizabilities of the materials involved. A number of approaches have been developed for estimating the value of the Hamaker constant; these can be broadly classified as microscopic and macroscopic (Visser, 1972; Nir, 1976). Microscopic theories attempt to relate the Hamaker constant to molecular properties, such as ionization potential, but these do not have a clear meaning for macromolecules such as proteins. An alternative, more rigorous framework for the estimation of this property is the macroscopic Lifshitz theory (Dzyaloshinskii et al., 1961). That the initial discussion of Lifshitz theory was in terms of quantum field theory placed it beyond the reach of much of the community interested in van der Waals interactions between macroscopic surfaces, but simpler derivations of the essential equations were presented subsequently (van Kampen et al., 1968), as were insightful guides to practical use of the equations (Gingell and Parsegian, 1972; Parsegian, 1975; Hough and White, 1980). The practical result that emerges from Lifshitz theory is that the van der Waals interaction energy can be estimated from the frequency-dependent dielectric spectra of the materials involved (including solvent) and the geometry. As discussed in more detail under Methods below, calculations for tractable geometries suggest that the material and geometric components are adequately represented by the Hamaker form (Eq. 1) at close separations (Mitchell and Ninham, 1972; Kiefer et al., 1976). Consequently, utilization of the Hamaker approach should be acceptable for arbitrary geometries, with the material property, in the form of the Hamaker constant, found from the dielectric spectra according to Lifshitz theory. Relatively scant dielectric/absorption data can be sufficient to calculate Hamaker constants of fair accuracy (tens of percent) (Parsegian and Ninham, 1969; Gingell and Parsegian, 1972; Hough and White, 1980), but it is apparent that detailed optical data should translate into a more accurate estimate of the Hamaker constant.

Thus modeling van der Waals forces in colloidal systems becomes a matter of approximating the geometry of the bodies and determining the value of the Hamaker constant. It is these two aspects that we investigate for proteins, with both the geometry and the material properties accounted for more thoroughly than has been done previously.

For many simple geometries, analytical expressions for the geometric part can be obtained; a large number of them are given in Lyklema (1991). A sphere is commonly used to approximate the shape of a protein, and for this case the analytical expressions for the van der Waals interaction energy are as follows (Hunter, 1986):

**Sphere-infinite plane:**

$$\Delta F_{\text{sphere/plane}} = -\frac{A_{132}}{6} \frac{R}{z} + \frac{R}{2R + z} + \ln \left( \frac{z}{2R + z} \right),$$  

(2)

where $R$ is the radius of the sphere and $z$ is the nearest distance between sphere and plane.

**Sphere-sphere:**

$$\Delta F_{\text{sphere/sphere}} = -\frac{A_{132}}{6} \left[ \frac{2R_1 R_2}{r^2 - (R_1 + R_2)^2} + \frac{2R_1 R_2}{r^2 - (R_1 - R_2)^2} \right]$$  

$$+ \ln \left( \frac{r^2 - (R_1 + R_2)^2}{r^2 - (R_1 - R_2)^2} \right),$$  

(3)

where $R_1$ and $R_2$ are the radii of the spheres and $r$ is the center-to-center distance. Eqs. 2 and 3 illustrate the short-
range nature of van der Waals interactions: because Hamaker constants are typically on the order of \( kT \), where \( k \) is the Boltzmann constant and \( T \) is the absolute temperature, the mathematical forms of the equations suggest that the free energy of interaction is significant only at separation distances much less than the sphere radii. The applicability of the sphere approximation at such short range is the first major aspect of the present work. We present calculations based on Eq. 1 as applied to geometries derived from protein crystallographic structures and show the results to differ appreciably from those given by Eqs. 2 and 3.

The second aspect of the work is the calculation, using Lifshitz theory and available spectroscopic information, of Hamaker constants for interactions of proteins with various other materials of interest. Although our results for protein-protein interactions here are in line with previous theoretical and experimental estimates, the overall conclusion of the work is that generally accepted estimates of the magnitude of van der Waals interactions involving proteins are much higher than the true values. This outcome has important implications for analyses of protein behavior in solution and near surfaces, even in the event that the discrepancies we calculate are offset by deviations from the assumptions inherent in our analysis.

**METHODS**

**Geometry**

We utilize the approach of Hamaker, but also incorporate the effect of protein molecular geometry. To accomplish this, the integral in Eq. 1 must be evaluated for geometries in which one or both interacting bodies have the shape of a protein molecule; we investigate both protein-protein and protein-plane interactions. In principle, the protein molecule can be represented as a single monolithic body. However, we find it more efficient computationally to perform the calculation as a volume integration over spheres representing each of the molecular groups comprising the macromolecule. A molecular group refers to any nonhydrogen atom (carbon, nitrogen, oxygen, etc.) and all hydrogens bonded to it (Bondi, 1968).

The use of molecular groups accounts for the different sizes of atoms or groups depending on their bonding state. To perform a piecewise volume integration in a consistent manner, one needs the piecewise volumes to sum to the correct total volume. Methods involving van der Waals radii are simplistic, as sums of the atomic volumes using van der Waals radii are consistently higher than volumes indicated by the molecular weight and specific volume of the proteins (Table 1). The Bondi group volumes represent only the volume occupied by the atoms comprising the groups; formation of a crystalline solid or a folded globular protein results in the packing of these groups with a density consistently close to 0.75 (Richards, 1974; Chothia, 1975). When the group volumes are rescaled accordingly, the total volume calculated with molecular groups agrees well with the expected values from bulk measurements for seven of the proteins that we studied (Table 2), for which data are tabulated by Creighton (1993).

For the van der Waals calculations, each molecular group was taken as a sphere with volume corresponding to the volume of the group with its center at the location indicated by the protein coordinate file, obtained from the Brookhaven Protein Data Bank (Bernstein et al., 1977). The van der Waals energy for a protein molecule and a surface was then calculated as the sum of the energies of each of the group spheres interacting with the surface, by application of Eq. 2. For a given protein and gap distance, the interaction energy was computed at 1669 evenly distributed orientations. Similarly, the van der Waals energy for two interacting protein molecules was calculated from Eq. 3 as the sum of the energies of each group sphere in one molecule with each in the other; 2500 evenly distributed orientations were used. Calculations were performed for nine proteins, namely those listed in Table 2 plus \( \alpha \)-amylase (Aspergillus oryzae) and pepsin (pig).

**Hamaker constants**

The essential physical picture of dispersion interactions that is formalized in Lifshitz theory is that the fluctuating dipoles within the interacting media establish electric fields that interact both constructively and destructively. The result of these fluctuating, many-bodied interactions is the formation of a standing wave between the two bodies whereby only certain modes, or frequencies, of electromagnetic radiation may pass. Although the allowed modes cannot be solved for explicitly, they do form an implicit dispersion relation \( D \) in terms of complex dielectric data on the imaginary axis. For two infinite planes, denoted by 1 and 2, separated by a distance \( z \) occupied by intervening medium 3, the expression for the dispersion interaction energy is given by

\[
\Delta F_{12}(z) = \frac{kT}{2\pi} \sum_{n=0}^{\infty} \int_0^\infty Q \, dQ \, \ln[D(i\varepsilon_n, Q)],
\]

where \( Q \) represents the wave vectors and the dispersion relation is given by

\[
D(i\varepsilon_n, Q) = 1 - \frac{y_1 y_2 e^{-Qz}}{y_1 + y_2}. \tag{5}
\]

**TABLE 1** Molecular volume of lysozyme calculated from its structure by means of several methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Volume (Å³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density/molecular weight*</td>
<td>16,700</td>
</tr>
<tr>
<td>PQMS program*</td>
<td>15,870</td>
</tr>
<tr>
<td>Sum of residue volumes²</td>
<td>17,900</td>
</tr>
<tr>
<td>Van der Waals radii¹</td>
<td>18,630</td>
</tr>
<tr>
<td>Van der Waals radii²</td>
<td>21,120</td>
</tr>
<tr>
<td>Van der Waals radii³</td>
<td>21,120</td>
</tr>
<tr>
<td>Group contributions²²</td>
<td>12,700</td>
</tr>
<tr>
<td>Group contributions/0.75</td>
<td>17,000</td>
</tr>
</tbody>
</table>

*Creighton (1993).
²Chothia (1975).
³Bondi (1964).
⁴Chothia (1975).
⁵McCammon et al. (1979).
⁶Bondi (1968).

**TABLE 2** Comparison of measured molecular volumes of proteins with computed volumes using molecular groups (Bondi, 1968)

<table>
<thead>
<tr>
<th>Protein</th>
<th>Measured volume (Å³/molecule)</th>
<th>Groups volume (Å³/molecule)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carbonic anhydrase (human)</td>
<td>34,800</td>
<td>34,900</td>
</tr>
<tr>
<td>Chymotrypsinogen A (bovine)</td>
<td>30,700</td>
<td>30,800</td>
</tr>
<tr>
<td>Myoglobin (sperm whale)</td>
<td>22,000</td>
<td>20,900</td>
</tr>
<tr>
<td>Ribonuclease (bovine)</td>
<td>15,000</td>
<td>16,100</td>
</tr>
<tr>
<td>Lysozyme (hen egg white)</td>
<td>16,700</td>
<td>17,000</td>
</tr>
<tr>
<td>Trypsin (bovine)</td>
<td>28,000</td>
<td>27,900</td>
</tr>
<tr>
<td>Cytochrome c (tuna heart)</td>
<td>14,600</td>
<td>13,800</td>
</tr>
</tbody>
</table>

Measured values are calculated from molecular weight and partial specific volume (data from Creighton, 1993).
in which

\[ y_{ab} = \epsilon_j(\omega) - \epsilon_{ij}(\omega) \]

(6)

with \( \epsilon_j(\omega) \) indicating the relative permittivity of medium \( j \) at the allowed complex frequencies \( \epsilon_{ij} \). The prime on the summation in Eq. 4 denotes the convention of dividing the \( n = 0 \) term by 2. By making a change of variables of the form \( \tau = 2\bar{Q} \), it is possible to combine Eqs. 4 to 6 into the same form as that given by Hamaker, i.e.,

\[ \Delta F_{12}(\tau) = -\frac{A_{132}}{2\pi^2} \]

(7)

with

\[ A_{132} = \frac{3kT}{2} \sum_{n=0}^{\infty} \sum_{\nu=1}^{\nu_s} \frac{(\nu_{13})^n}{(\nu_{12})^n}; \]

(8)

the integral in Eq. 4 has been expanded in a power series and integrated term by term.

Although the above result, based on previous developments by Hough and White (1980) and Hunter (1986), was derived only for the case of two infinite planes, it is such that the interaction is characterized by the Hamaker constant, and the remainder of Eq. 7 is geometric. Similarly, calculations for the only other tractable geometry, sphere-sphere, suggest that the Hamaker formulation (Eq. 1) remains valid (Mitchell and Ninham, 1972; Kiefer et al., 1976). Consequently, utilization of the Hamaker approach is assumed valid for arbitrary geometries, with the Hamaker constant found from Eq. 8.

Evaluation of \( A_{132} \) is facilitated by the fact that the complex dielectric constant appearing in Eq. 8 via Eq. 6 has a relatively simple physical basis and furthermore can be constructed from a small amount of experimental data (Ninham and Parsegian, 1970; Hough and White, 1980). If the permittivity is expressed in terms of its real and imaginary components, i.e.,

\[ \epsilon(\omega) = \epsilon'(\omega) + i\epsilon''(\omega), \]

(9)

then \( \epsilon' \) corresponds to a phase lag in the dielectric response and hence to energy dissipation. This dissipation is observable in the absorption spectrum of the material. Furthermore, \( \epsilon' \) is related to the desired \( \epsilon(\bar{Q}) \) by a Kramers-Kronig relation (Landau and Lifshitz, 1960):

\[ \epsilon(\bar{Q}) = 1 + \frac{2}{\pi} \int_0^\infty \frac{x \epsilon''(x)}{x^2 + \bar{Q}^2} \, dx. \]

(10)

Additionally, for frequencies at which no absorption takes place, i.e., \( \epsilon'' = 0 \), the complex dielectric permittivity consists only of the real portion \( \epsilon' \) and is given by the square of the refractive index of the material. Using this type of simple experimental information, others have computed Hamaker constants and interaction energies for a number of planar and layered materials; some summaries are included in Visser (1972), Hough and White (1980), Lyklema (1991), and Israelachvili (1992).

Characterization of the complete absorption spectrum thus becomes the primary task in determining the Hamaker constant. For many materials, a simple construction involving only the static dielectric constant, refractive index in the visible, and an ultraviolet relaxation characterized by a Cauchy plot provides sufficient information to estimate the Hamaker constant reasonably well (Hough and White, 1980). Although this method represents an accessible first approximation, clearly it is desirable to test the validity of the method against Hamaker constants calculated from the complete absorption spectrum. Data of this detail are available for a limited number of materials, obtained by means of reflectance or transmission experiments over a wide range of frequencies.

For convenience of use in Eq. 8, the dielectric data are generally translated into a mathematical formula. The relaxations that result in the decay from the static dielectric constant at zero frequency to a value of 1 at very high frequencies are assumed to follow the forms of a Debye relaxation in the microwave, damped oscillator in the infrared and ultraviolet, and plasma relaxation in the x-ray (Parsegian, 1975). On the whole, this approach was followed here as well; however, it has been argued that the plasma resonance is implicitly included in the ultraviolet term (Hough and White, 1980), and so it was not included here. Relaxations in the microwave occur for polar materials of very high dielectric constant, such as water, and are fitted to a form that corresponds to dipole reorientation (Parsegian, 1975). For most materials, the most significant relaxations occur in the infrared and especially the ultraviolet, where the damped oscillator form is used. The complete form used for the decay is then

\[ \epsilon(\omega) = 1 + \sum_{k} \frac{d_k}{1 - i\omega \tau_k} + \sum_{j} \frac{f_j}{\omega_j^2 + g_j(-i\omega) + (-i\omega)^2}, \]

(11)

where \( \epsilon(\omega) \) is the dielectric response at frequency \( \omega \), \( 1/\tau_k \) is the microwave relaxation frequency, \( d_k \) is the microwave decay amplitude, \( \omega_j \) is the oscillator frequency (infrared or ultraviolet), \( g_j \) is the bandwidth of the relaxation, \( f_j \) is the oscillator decay amplitude, and \( \omega = \sqrt{-T} \).

The transmission data available for many compounds cover only the range of frequencies from the visible through the far ultraviolet. The real and complex parts of the dielectric permittivity of substances for which data were available were fitted to a damped oscillator form (third term in Eq. 11) using a Levenberg-Marquardt method with bounds (nonnegativity) on the variables, implemented via the IMSL routine DBCLSI (Visual Numerics, Houston, TX).

For interactions involving proteins, dielectric data for bovine serum albumin (BSA) were used (Inagaki et al., 1975). This data set was fitted to three damped oscillators in the ultraviolet. One infrared term was also included, corresponding to a characteristic absorption between the amide I and amide II bands at 1530 and 1650 cm\(^{-1}\), respectively. A value of 3 was used for the static dielectric constant. Likewise, fits were made to polystyrene and water dielectric data, which are also available (Heiler et al., 1974; Parsegian, 1975). For quartz and many of the polymers, no complete dielectric data were available; for these materials, the construction of Hough and White (1980) was used. The characteristic UV absorption frequency is usually obtained in this construction from a Causky plot of refractive index as a function of frequency in the visible region of the spectrum. For an absorption that can be approximated as having a sharp peak (negligible bandwidth) and assuming negligible absorption in the visible region over which the refractive index is measured, the Causky plot assumes the form (Hough and White, 1980)

\[ n^2 - 1 = (n^2 - 1) \frac{\omega^2}{\omega_{UV}^2} + C_{UV}, \]

(12)

in which \( n \) is the index of refraction of the material, \( \omega_{UV} \) is the characteristic ultraviolet frequency, and \( C_{UV} \) is the peak amplitude. The peak amplitude \( C_{UV} \) is related to the Ninham-Parsegian representation (Eq. 11) by

\[ C_{UV} = \frac{f_{UV}}{\omega_{UV}}. \]

(13)

The resulting functions in the form of Eq. 11 were used in Eqs. 6 to 8 to compute the Hamaker constant. The inner series in Eq. 8 was computed to 500 terms, and the outer series to 2000 terms, both of which were found to be more than sufficient. Results were checked against published values (Hough and White, 1980) for identical inputs, and agreement to within 1% was obtained. Calculations were performed on an IBM RISC 6000 model 590 computer and required only a few seconds to obtain the Hamaker constant for any set of input data.
RESULTS

Geometry

As discussed earlier, van der Waals interactions are of very short range relative to the dimensions of protein molecules: consequently, we look first at results for protein-surface and protein-protein nearest separation distances of 1 Å. For such gaps the applicability of continuum theories is clearly in question, but this arbitrary value facilitates illustration of the effects involved, and we later consider the effect of gap distance explicitly. For the relatively small and well-characterized proteins that we studied, frequency histograms showing the orientational distributions of the interaction energy at 1 Å separation distance are quite similar to one another: sample results for lysozyme and chymotrypsinogen A are shown in Fig. 1. The frequency histograms for the protein-surface interaction energies at 1 Å (based on results for nine proteins) can be represented adequately by log-normal distributions with σ values of 0.143 ± 0.015, with no apparent correlation with protein size. For protein-protein interactions the tails are typically longer, with a small number of highly favorable orientations observed, so the fit to a continuous distribution is less successful.

In addition to its failure to capture the existence of a distribution of energies with respect to orientation, a sphere model provides a poor estimate of the magnitude of the van der Waals interaction energy and does not accurately represent the scaling with respect to size or molecular weight. Fig. 2a shows the average van der Waals interaction energies for several globular proteins interacting at 1 Å nearest separation from a planar surface; the averages are arithmetic ones computed over orientational distributions of the kind shown in Fig. 1. Fig. 2a also shows corresponding results for spheres of equivalent volumes, calculated using Eq. 2. Correspondingly, Fig. 2b shows the van der Waals interaction energies of two proteins or two spheres at a gap distance of 1 Å. Clearly, the assumption of a spherical geometry overestimates the magnitude of the average van der Waals interaction energy. In fact, for the protein-surface case, none of the 1669 uniformly distributed orientations is

![Figure 1](image-url)

**FIGURE 1.** Orientational distribution of dispersion interaction energies, scaled relative to the Hamaker constant, for lysozyme and chymotrypsinogen A at 1 Å from a planar surface or an identical molecule. Solid lines indicate fits to log-normal distribution. Protein-surface: (a) lysozyme; (b) chymotrypsinogen. Protein-protein: (c) lysozyme; (d) chymotrypsinogen.
as attractive as the equivalent sphere (see Fig. 2 a). For two protein molecules, only one orientation in 2500, and then for only one of the proteins examined, produces an interaction energy greater in magnitude than that of two spheres (see Fig. 2 b). Furthermore, these comparisons are based on intervening gap distances; the discrepancies would, in general, be even larger for calculations based on the centers of the protein molecules, in view of the relative compactness of spheres of equivalent volume.

The correlation of the average protein-surface van der Waals energy with respect to molecular volume (Fig. 2 a) suggests that a universal curve can be used to approximate the magnitude of protein-surface interactions at 1-Å separation:

\[
\frac{\Delta F}{A} \text{ave} = 0.40 \text{MW}^{0.20} \quad (R^2 = 0.864),
\]

where \(\Delta F\) is the dispersion energy, \(A\) is the Hamaker constant, \(\text{MW}\) is the molecular weight of the protein, and \(R^2\) is the correlation coefficient. The corresponding relationship for the protein-protein case is

\[
\frac{\Delta F}{A} \text{ave} = 0.15 \text{MW}^{0.17} \quad (R^2 = 0.718).
\]

Thus the scaling of interaction energy with molecular weight is weaker than that for a sphere—only about a one-fifth or one-sixth power dependence for the protein versus roughly one-third for a sphere—and the magnitudes of the interactions are generally much weaker than would be expected from a spherical representation.

The results presented above for a 1-Å gap can be used as the basis for more general calculations of van der Waals interactions involving proteins. These interactions are short-ranged: Eqs. 2 and 3 can be used to show the distance dependence of the van der Waals interaction energy for both sphere-surface and sphere-sphere interactions to be roughly \(1/z\), where \(z\) is the gap distance, for moderately small separations. In Fig. 3 this dependence is compared with the average value for lysozyme molecules; the shaded regions represent the range of energies covered by the different orientations studied. The detailed shape effects are seen to be less important for large \(z\); the average interaction energies approach the sphere values, and the range of energies becomes correspondingly smaller in each case. For smaller \(z\), however, the shapes of the interacting surfaces become more influential, and the divergence between the protein average and the sphere results increases as the intervening distance drops. In both the protein-surface and protein-protein cases, however, the most favorable interactions remain fairly close to the sphere results throughout, at least on the logarithmic scale used. For protein-protein interactions, the average shows a pronounced shift from the upper to the lower side of the range as the gap distance becomes smaller, reflecting the low frequency of highly favorable orientations (see Fig. 1), an effect not apparent for the protein-surface case.

Hamaker constants

Although Lifshitz theory provides a rigorous basis for calculating Hamaker constants within a continuum framework, the values obtained depend strongly on the accuracy and the amount of detail in the representation used of the spectral data. We have used both detailed representations of spectral data and the Cauchy plot method (Hough and White, 1980) in this work. For proteins, both the Cauchy plot as well as extended data from transmission spectroscopy indicate that the principal relaxations occur outside the range accessible by analytical spectroscopy. We have constructed Cauchy plots from refractive index data for several proteins (McMeekin et al., 1964), and they all indicate a characteristic UV absorption wavelength of about 125 nm, corresponding to 9.9 eV. The complete spectral data for BSA (Inagaki et al., 1975) show two distinct peaks and a shoulder, which were fitted to the parameters shown in Table 3; the Cauchy...
frequency would appear to represent a reasonable average of the whole spectrum.

Applying these results to the computation of Hamaker constants for protein-protein interactions in air yields similar results for the two methods: 23.42 $kT$ for the detailed representation and 21.11 $kT$ for the Cauchy plot. Thus the two methods differ by about 10%. For comparison, the interaction of two films of polystyrene across air serves as a useful example of the potential discrepancies between the use of complete spectroscopic data and the Cauchy plot method. The former method yields a Hamaker constant of 20.3 $kT$, based on the experimental data of Inagaki et al. (1977) and fitting parameters of Parsegian (1975), whereas the latter method (Hough and White, 1980) gives 15.7 $kT$. Results for other systems confirm that discrepancies on the order of tens of percent are typical between results obtained by the two methods.

Because the dispersion interaction is a consequence of differences (i.e., not just absolute values) in polarizability across the electromagnetic spectrum, the accuracy of a construction in a region of the spectrum can be more or less important, depending on the other materials involved in the interaction. This feature is encountered most often when Hamaker constants are calculated for materials separated not by air, but by a dielectric medium, most commonly water. Accurate representation of the spectral information for this intervening medium is thus imperative. For the original water data (Heller et al., 1974), we have used an improved set of fitting parameters (Roth and Lennhoff, 1996) in preference to those of Parsegian (1975), and obtained protein-water-protein Hamaker constants of 3.10 $kT$ and 2.68 $kT$ by the detailed and Cauchy plot methods, respectively. Thus the strength of the interactions is greatly attenuated by the presence of the water, whereas the relative discrepancy between the predictions of the two methods is slightly greater.

The Hamaker constant values that we calculate are a little higher than the range suggested by Parsegian and Brenner (1976), when accounting is made of the fact that their calculations do not include the zero-frequency contribution to the dispersion energy. This component is screened in the presence of electrolyte, although the extent of screening is somewhat of a matter of debate (Mahanty and Ninham, 1976). The zero-frequency contribution to the Hamaker constant depends only on the differences in static dielectric constant between adjoining materials; for low-dielectric materials interacting across water, this contribution is roughly constant at a value we calculate to be approximately 0.75 $kT$. The discrepancy between the values computed by us and by Parsegian and Brenner probably is a consequence of the details of the fit of the dielectric data, as discussed above for water.

Although protein-protein Hamaker constants are important for solution thermodynamic properties, the contribution of van der Waals interactions to protein adsorption is better characterized by Hamaker constants for protein interactions with typical sorbents. Hamaker constants for the interaction of BSA with a variety of sorbents, with water as the intervening medium, are presented in Table 4. Although the absolute values are probably accurate only to about 30%, the variation among materials is large enough to indicate the relative strength of dispersion interactions in each of the systems.

### DISCUSSION

The results presented above have significance for both observed biophysical behavior of proteins and for compu-
tional approaches for modeling proteins. However, in evaluating this significance it is important also to assess
the validity of the approach used and hence of the results.
Central to this issue is the fact that we have applied a theory
that assumes rigid, homogeneous, isotropic media for sys-
tems that may display considerable structural variability
with position, direction, and time. In addition, the charac-
teristic length scales in our calculations cover a range that
tends down to the dimensions of individual water mole-
cules, whereas the theory assumes continuum properties.
Thus the results should be used with some caution, but
nonetheless we believe that their essential features ade-
quately describe the behavior of many protein systems.

The results of our geometric computations indicate a
major departure from generally accepted values for protein
der Waals interactions. Both the extent of the orienta-
tional dependence (Fig. 1) and the disparities between the
average energies and those determined using the sphere
approximation (Fig. 2) are due to what may be classified as
surface roughness. For a folded protein of rigid conforma-
tion, only small, high-curvature parts of the apposing sur-
faces are able to approach each other closely, and much of
the molecular surface is excluded to either a planar surface
or to another protein molecule in any given configuration.
In the case of protein-protein interactions, situations in which
relatively extended parts of apposing surfaces can come into
contact lead to highly attractive interaction energies due to
der Waals interactions as well as additional contribu-
tions such as solvation and hydrogen bonding. That such
gemetric complementarity is essential for specific binding
has been widely recognized (e.g., Hendrickson et al., 1987),
as has the importance of the correct orientation for mole-
cules adding to a protein crystal. Less attention has been
paid to phenomena involving less specific interactions, such
as aggregation and precipitation. Classical colloidal theories
(see, e.g., Hunter, 1986; Lyklema, 1991; Israelachvili,
1992) usually describe aggregation as resulting when van
der Waals attraction overcomes electrostatic repulsion be-
tween like particles. The very small number of favorable
orientations and low average interaction energy, at least for
van der Waals interactions, indicates that aggregation is less
likely for proteins than for equivalent spheres; the surface
roughness may thus contribute to a form of steric stabiliza-
tion of the protein molecules in solution. Furthermore, the

persistance of the orientational dependence at longer range
may affect the dynamics of processes such as aggregation,
precipitation, crystal growth, and ligand binding.

Although the constraint of rigidity is typical of colloidal
systems, it is not always satisfied for proteins. Because the
dispersion force between a protein and a sorbent depends
strongly on their geometric complementarity, it can be sur-
mised that an adsorbing protein molecule might gain a
substantial energetic reduction by adapting its shape to that
of the interface. This might explain partially the tendency
for proteins to denature at surfaces (MacRitchie, 1976;
Kondo et al., 1992; Smith and Clark, 1992; Norde and
Favier, 1992). Likewise, the sorbent could adapt to the
adsorbing protein molecule, as is believed to be the case for
polymer "tentacle" ion-exchange adsorbents (Müller, 1990).
Structural changes that enhance complementarity in protein-
protein interactions have been studied more extensively,
e.g., by examining changes in the crystallographic struc-
tures of molecules when bound in a complex (e.g., Davies et
al., 1988). In addition to such induced structural changes,
conformational fluctuations are also likely to affect the
nature of interactions, including the separation distances
between different parts of the apposing surfaces. Clearly,
the approach used here is not capable of accounting directly
for such effects, which presumably reflect dynamic events
at very short range.

Direct experimental verification of the results is difficult
because a variety of forces are manifested in any experi-
mental system, and few of them are well enough understood
quantitatively in proteins for the dispersion contribution to
be identified unambiguously. However, our protein-water-
protein Hamaker constant calculations can be compared
with measured values obtained on relatively extended inter-
acting surfaces, where geometric effects are less important.
Most of the results are from surface forces apparatus exper-
iments, the most extensive of which are those of Afshar-Rad
et al. (1987), who estimated the Hamaker constant for
protein films across water to be in the range $1.0 - 2.2 \times
10^{-20} \text{J} (10^{-20} \text{J} = 2.4 \text{kT at } 20°C)$; our value of $3.1 \text{kT thus}
falls near the middle of this range. In a more recent study
(Helm et al., 1991), a value of $4 \times 10^{-21} \text{J}$ was determined
for the interaction of films of streptavidin in a lipid bilayer
across an aqueous 1 mM NaCl solution. The authors attrib-
uted the lower Hamaker constant to the incorporation of the
receptor in a bilayer film that itself possesses a Hamaker
constant of $1.5 - 7.0 \times 10^{-21} \text{J}$, but other possibilities sug-
gested by this work are the screening of the electrolyte and
a geometric reduction due to the fact that their films con-
sisted of a low (~5%) surface coverage of protruding pro-
tein molecules on an otherwise planar lipid bilayer. Com-
parison can also be made with the data of Srivastava (1966),
who estimated a Hamaker constant of $1.8 \times 10^{-20} \text{J}$ from
the aggregation kinetics of BSA emulsion droplets; again,
this value is quite close to that which we calculate.

Thus our calculated values of Hamaker constants appear
reasonable. However, the circumstances under which they
can be used to characterize protein van der Waals interac-

<table>
<thead>
<tr>
<th>Material</th>
<th>$A_{BSA} (\text{kT})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSA</td>
<td>3.10</td>
</tr>
<tr>
<td>Polyvinyl chloride</td>
<td>2.29</td>
</tr>
<tr>
<td>Poly(methylmethacrylate)</td>
<td>1.96</td>
</tr>
<tr>
<td>Quartz</td>
<td>1.65</td>
</tr>
<tr>
<td>PS</td>
<td>1.54</td>
</tr>
<tr>
<td>Poly(tetrafluoroethylene)</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Data for polystyrene, polyvinyl chloride, poly(methylmethacrylate), and poly(tetrafluoroethylene) from Hough and White (1980).
tions over length scales much smaller than the protein dimensions remains unclear. The central considerations here are those of the heterogeneity and anisotropy of protein molecules, and the question of what minimum gap distance is needed for the continuum assumption to be applied to the intervening aqueous medium. Clearly, length scales significantly greater than the size of a water molecule are appropriate. Simulations suggest that the water structure around a protein or near a large particle approaches bulk properties within two or three molecular layers (e.g., Torrie and Patey, 1991; Gerstein and Lynden-Bell, 1993). On the other hand, surface force measurements between mica sheets, summarized by Israelachvili (1992), show details of the water molecular layering below about 15 Å gap spaces; however, these details are no longer apparent for surfaces that are rougher and/or more fluid than mica. Furthermore, a compensating factor for proteins is that at a given minimum gap distance in a protein system, only very small fractions of the apposing surfaces are, in fact, in such close proximity. It appears reasonable, therefore, to use our quantitative results down to gap distances on the order of 10 Å. For relatively small globular proteins, the interaction here is very weak, but the semiquantitative conclusions of the geometric computations (guided by Eqs. 14 and 15 along with the form of Fig. 3) should remain largely valid, even at smaller gap distances in the absence of appreciable structural changes. The quantitative uncertainties obviously increase at smaller gap distances, though, especially in the values of the Hamaker constants. These observations may be compared with the application of continuum molecular electrostatics computations to the interpretation of intermolecular and intramolecular processes that involve short-range interactions such as enzyme-substrate binding, protein adsorption, and protein pKₐ shifts (Bashford and Karplus, 1991; Yoon and Lenhoff, 1992; Roush et al., 1994).

Even within the constraints of these caveats, the computed Hamaker constants can help shed light on observed behavior. Specifically, the results in Table 4 are relevant to analysis of the adsorption of proteins to solid interfaces, which is involved in many physiological phenomena and in processes of medical and engineering consequence. Different views exist as to the primary mechanisms underlying protein adsorption, and these are likely to vary strongly with the particular protein, sorbent, and environmental conditions. Our calculations and experimental evidence (Roth and Lenhoff, 1995) indicate that van der Waals interactions, although generally not strong enough to provide the total driving force for many of the high-affinity isotherms that are observed, contribute enough to adsorption to affect the trends observed with respect to protein size. The strength of this interaction is indicated by the Hamaker constant for a protein interacting with the surface through a solvent (generally water or aqueous electrolyte), values of which are shown in Table 4. Notwithstanding the uncertainties discussed above, the variation among materials is large enough to indicate the relative strength of dispersion interactions in each of the systems. For example, the inert nature of poly-

(tetrafluoroethylene) is reflected in its unusually low Hamaker constant. On the other hand, proteins would be expected to have a strong interaction with quartz and polymers such as polystyrene. In fact, adsorption of proteins has been found to occur on polystyrene under conditions of electrostatic repulsion as well as electrostatic attraction (Arai and Norde, 1990; Haynes et al., 1994), although there are certainly a number of other factors that may also be at play in this process.

Combining the geometric properties and the Hamaker constant provides a convenient route to the description of protein-surface and protein-protein interactions; it is particularly for the latter application that the colloidal approach is an expedient alternative to the more widely used atomistic methods. An illustration of their relative behavior is provided by a comparison of the van der Waals interaction energies computed with the Hamaker method and those calculated based on OPLS (optimized potentials for liquid simulation) parameters (Jorgensen and Tirado-Rives, 1988). The interaction energy calculated is for a system with a high degree of surface complementarity, viz. that between the antibody D1.3 Fab fragment (Amit et al., 1986) and its antigen, lysozyme, as the molecules are moved apart along a line perpendicular to their approximate contact surface. The Hamaker constant used is that for protein-air-protein interaction to compare to the OPLS parameters, which have been optimized for short-range interactions; we have omitted water molecules from the calculation.

Results are shown in Fig. 4 for gap distances of 5 Å and greater relative to the crystallographic structure; results for smaller gaps are not shown because the van der Waals radii overlap for the crystal structure of the complex and for the first several angstroms when the complex is pulled apart. Both techniques predict a highly attractive interaction energy much greater than kT at a distance of 5 Å, whereas at larger gap distances the OPLS result is quite consistently about double that of the colloidal method. The distance

![FIGURE 4 Comparison of colloidal (solid line) and OPLS (broken line) calculations of dispersion interaction energies between a lysozyme molecule and IgG D1.3 Fab fragment. Distance is relative to the crystallographic coordinates on an axis perpendicular to the approximate surface of contact.](image-url)
dependence for this interaction is approximately $1/z$ at long range as before, but at short range it is much stronger, in view of the complementarity of the interacting surfaces. In general the $1/z$ dependence is expected to prevail at long range, whereas the short-range dependence will vary depending on the degree of complementarity present. Although the two methods yield essentially the same distance dependence, the OPLS parameters yield interaction energies that are greater in magnitude. This is presumably because they are fitted to small-molecule physical property data and include implicitly all attractive interactions, such as hydrogen bonding. Taking this into consideration, the colloidal approach appears to provide an efficient route to estimation of van der Waals interaction energies, except at very short range. It may thus be a useful tool to incorporate into multimolecular calculations, such as Brownian dynamics simulations (Zhou, 1993).

**CONCLUSIONS**

We have calculated the geometric and material contributions to van der Waals interaction energies involving proteins. Our approach and results have two major implications. First, they provide insight into the role of van der Waals interactions in protein solutions, particularly through the findings that the geometric factor is considerably smaller than is estimated using idealized shapes such as spheres, and that its dependence on protein size is much weaker. Because of these characteristics, the use of idealized models may severely overestimate the magnitude of dispersion interactions, except when appreciable complementarity of approaching surfaces is either present initially or is induced by the interaction. These results are pertinent to protein properties such as solubility and adsorption behavior. The second implication of the study is that the colloidal approach provides a useful approach to capturing the same types of structural detail as atomistic methods, but in a more efficient fashion. This approach is likely to be especially useful in developing rational simplifications for use in multimolecular computations.

We acknowledge useful discussions with V. A. Parsegian and D. B. Weislauf.

This work was supported by grants from the National Science Foundation (CTS-9111604 and BCS-9210401) and the National Aeronautics and Space Administration (NAGW-2798).

**REFERENCES**


