Gas-Phase Spectroscopy of Biomolecular Building Blocks

Mattanjah S. de Vries\(^1\) and Pavel Hobza\(^2\)

\(^1\)Department of Chemistry and Biochemistry, University of California, Santa Barbara, California 93106; email: devries@chem.ucsb.edu
\(^2\)Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, 166 10 Prague 6, Czech Republic; email: pavel.hobza@uochb.cas.cz

**Key Words**
biomolecules, REMPI, computational chemistry, spectral hole burning, jet cooling

**Abstract**
Gas-phase spectroscopy lends itself ideally to the study of isolated molecules and provides important data for comparison with theory. In recent years, we have seen enormous progress in the study of biomolecular building blocks in the gas phase. The motivation for such work is threefold: (a) It is important to distinguish between intrinsic molecular properties and properties that result from the biological environment. (b) Gas-phase spectroscopy of clusters provides insights into fundamental interactions and into microsolvation. (c) Gas-phase data support quantum-chemical calculations. This review focuses on the current status of (poly)amino acids and DNA bases. Recent results help elucidate structure and hydrogen-bonded interactions, as well as showcase a successful interplay between theory and experiment.
1. INTRODUCTION

Research in the past decade has shown considerable progress in the study of biomolecular building blocks in the gas phase, which makes it possible to distinguish between intrinsic molecular properties and properties resulting from the biological environment. The details of fundamental biochemical characteristics and interactions are often hidden by macroscopic solvent effects, interactions with other molecules, constraints imposed by the macromolecular backbone, and averaging. Thus studying individual biomolecules, free of the external biological environment, reveals the intrinsic properties of the most basic biomolecular processes. Detailed processes are further elucidated by observing how the intrinsic properties of the isolated molecules change with biomolecular interaction and with the sequential addition of single-solvent molecules. For example, isolated DNA bases can form hydrogen-bonded pairs in many arrangements, but study of isolated guanine-cytosine pairs found the Watson-Crick structure to be distinguished by unique photochemistry, and isolated model peptides show folding motifs without solvent stabilization in very short sequences.

The study of isolated molecules is the prime domain of gas-phase research, which also provides the best data for comparison with theory. Until recently, DNA bases, amino acids, and related molecules could not be studied in the gas phase for practical reasons. Experimentally, these compounds have low vapor pressures, and they decompose when heated. Computationally these systems were intractable on a high theoretical level because of their size and because correlation or dispersion energy would have to be included to properly describe them. Therefore, semiempirical quantum-chemical methods and even ab initio Hartree-Fock and density-functional theory (DFT) methods were of limited use. In both areas, significant progress has now been made.

In recent years the problem of transferring biomolecules into the gas phase intact has largely been solved, making a rich tool chest of gas-phase techniques available for the study of these molecules. This review focuses on the spectroscopy of neutral biomolecular building blocks. Although optical spectroscopy can provide detailed structural information, it also poses additional practical challenges. High resolution requires low internal temperatures; furthermore, densities in the gas phase are usually quite low. Therefore, among the most ubiquitous approaches are those that employ jet cooling in combination with electronic spectroscopy. The latter allows for the detection of ions from multiphoton ionization or fluorescence to overcome sensitivity limitations. Other spectral ranges, such as the infrared (IR), can then be studied by double-resonance techniques. We thus mainly devote this review to covering these types of experiments.

Spectroscopy requires computational theory for the interpretation of spectra, but the increasing size of the molecules studied challenges state-of-the-art computations. New algorithms and force fields need to be developed, which in turn require gas-phase experiments for calibration. In fact, simultaneous progress in theory and experiment has driven the explosive progress in this field.

We mainly distinguish between two types of biomolecular building blocks for the purpose of this review:
1. (Poly)amino acids. For these compounds, the emphasis of gas-phase spectroscopy is on the study of conformations. Line shifts in combination with high-level calculations can help map conformational landscapes. Shape is crucial in biology, and the hope is that by recognizing fundamental propensities, it will eventually be possible to add to our understanding of protein folding.

2. Nucleobases. For these compounds the emphasis is on elucidating their photochemistry and their hydrogen (H)-bonding interactions, which shed light on base-pairing dynamics and on the details of solvation.

Other types of biomolecular building blocks studied in the gas phase include sugars (1), and smaller molecules and their derivatives that constitute the chromophores of the major building blocks, such as purines and pyrimidines as nucleobase chromophores, and indole as an amino-acid chromophore.

2. EXPERIMENTAL TECHNIQUES

2.1. Laser-Desorption Jet Cooling

Following early reports on the laser desorption of large molecules (2, 3), Cable et al. (4) studied the spectroscopy of tryptophan di- and tripeptides by resonance-enhanced multiphoton ionization (REMPI), as well as by laser-induced fluorescence (LIF), following laser desorption. Li & Lubman (5) demonstrated that for laser-desorbed tyrosine and related structural analogs, electronic spectroscopy can be a sensitive probe of small structural changes in related biological compounds. Arrowsmith et al. (6) and Meijer et al. (7) explored the characteristics of the technique.

The process of neutral laser desorption is poorly understood. A few mechanisms have been proposed (8, 9), and only a small number of experiments has been reported to study the process (10–12). In nonresonant ionization (e.g., at 193 nm or 266 nm), one can control fragmentation by the proper choice of wavelength and laser fluence. Tembreull & Lubman (13) and Grotemeyer et al. (14) have shown that fragmentation can be controlled by the ionization laser fluence. This control can be used to the advantage of structural analysis. Higher ionization efficiencies have been obtained by femtosecond laser pulse excitation, involving quasi-resonant effects (15).

2.2. Spectroscopy

Electronic spectroscopy in molecular beams includes LIF and REMPI (see Figure 1). The latter has the advantage of mass selectivity (16). Isomer distinction is possible with double-resonance techniques, which appear schematically in Figure 1. In these techniques, also called spectral hole burning (SHB), a burn laser pulse resonantly depletes the ground state and is followed approximately 100 ns later by a probe laser pulse, which is tuned to a vibronic transition of a single specific isomer. The resulting decrease in the probe laser signal thus provides isomer-specific (and in the case of a REMPI probe, also mass-selective) spectroscopy. Ultraviolet-ultraviolet (UV-UV) double resonance allows the deconvolution of a vibronic spectrum in the
contributions from different isomers, whereas IR-UV double resonance allows for structural assignments (17).

The red shifts of the NH-, OH-, C=O-stretching vibrations and the blue shifts of OH and NH in-plane bending vibrations upon hydrogen bonding often allow an unambiguous assignment of H-bonded structures (18, 19). Therefore the wavelength range of 2800–3600 cm\(^{-1}\) provides an excellent structural tool. This range is available from table optical parametric oscillator/amplifier systems. Recent extensions of tabletop laser systems to include the C=O stretch frequency at approximately 1800 cm\(^{-1}\) have enhanced this capability (20). Experiments at free-electron laser facilities cover lower frequencies (1, 19). Such lower-frequency experiments are in principle more diagnostic for details of side-group conformations and can serve as tests for computations that include anharmonic contributions to the potentials. However, the resolution of such lasers is generally limited to approximately 1% of the wavelength.

The isomer selectivity distinguishes molecular-beam spectroscopy from cold-matrix spectroscopy together with the absence of matrix effects. Typical UV resolutions of the order of a wave number combine with internal temperatures of the order of 10–20 K (7). The use of pulsed-dye amplifiers provides higher resolutions to allow for rotationally resolved spectra (21). In the near-IR range, optical parametric oscillator devices achieve resolutions on the order of a few wave numbers, whereas free-electron lasers in the mid-IR typically provide 10% resolution of the frequency.

We briefly describe related techniques that fall outside the scope of this review below. Stimulated emission pumping–hole filling spectroscopy and stimulated emission pumping–induced population-transfer spectroscopy provide insights into isomerization pathways (22). Microwave spectroscopy can lead to the accurate structural determination of small molecules (23). Helium droplets are unique because of the low temperature; among the possible applications is polarization-dependent IR spectroscopy (24). Populations of isomers in helium droplets may differ from those in supersonic beams because the timescale of cooling is much faster in the former. Another possible difference in the resulting spectroscopy stems from the possibility of
the He environment causing matrix effects. Finally, all results included in this review refer to the spectroscopy of neutral molecules. Limited work exists thus far on the optical spectroscopy of ions, although many other forms of ion spectroscopy exist, such as ion-drift spectroscopy and unimolecular reactions.

3. THEORY

3.1. Molecular Structure and Geometry

Structure and properties of molecular clusters are mainly determined by noncovalent interactions. A theoretical description of these interactions is difficult and requires the most accurate methods of computational chemistry available. Two structural motifs dominate biomolecular building blocks and their clusters: H-bonding and stacking. For the latter, London dispersion energy is the main source of stabilization. Whereas H-bonding interactions are well understood, stacking interactions are difficult to treat computationally. Figure 2 illustrates a case study for phenylalanine-glycine-glycine (Phe-Gly-Gly) for which double-resonance spectroscopy provides vibrational spectra of four conformations, one of which (Figure 2b) is stabilized by dispersive forces (25). The challenge for theory is to find the lowest-energy conformations, which can be tested by comparing their vibrational frequencies with the experimental spectra. Figure 2 shows the four calculated conformations and the corresponding comparison of frequencies. The following sections discuss the computational strategies for dealing with a conformational phase space of this size, and possibly larger.

3.1.1. Strategy. The aim of theoretical studies is the determination of minimum-energy structures, as well as their populations and properties at experimental conditions. Although it is sufficient to merely study the potential-energy surface (PES) for a system at 0 K, a complete description at nonzero temperatures requires extension to the free-energy surface (FES). The PES and FES differ; for example, the global minimum at both surfaces might correspond to completely different structures. The PES of biomolecular clusters is rich and contains a large number of energy minima. Finding the global minimum requires an efficient searching procedure. Molecular dynamics (MD) simulations, together with quenching (MD/Q), can provide a full description of the PES. However, generally MD/Q calculations employ empirical potentials, limiting their applicability and accuracy. The empirical potential used should be carefully a priori and a posteriori tested. Inevitably the combination of inaccurate potentials and high-level MD simulations yields inaccurate results. Therefore, structures obtained with the MD/Q treatment should be reoptimized at a high correlated ab initio level. Subsequently, detailed properties, such as vibrational frequencies, can be determined for the most stable structures of the PES or FES with a high-level quantum-chemical treatment.

3.1.2. Structure and geometry. Standard gradient optimization performed at the correlated Møller-Plesset second-order perturbation level (MP2) with a basis set of TZ + P quality [e.g., Dunning’s cc-pVTZ (correlation-consistent polarized
Figure 2

Optimized minimum energy structures for phenylalanine-glycine-glycine (Phe-Gly-Gly) (right-hand column). (b–e) Computed frequencies as stick spectra and infrared-ultraviolet spectral hole burning (IR-UV SHB) spectra, recorded with the probe laser at the wavelengths shown by the arrows in corresponding colors in the REMPI spectrum (a).

Valence triple-zeta basis set] generally yields sufficiently accurate results for H-bonded, stacked, T-shaped, and other structures of molecular clusters (26, 27). Further improvement is possible with counterpoise-corrected gradient optimization, which is, however, considerably more time consuming (28, 29). Resulting structures are close to the experimental ones, and thus may have predictive value in the absence of
Stabilization energies strongly depend on the theoretical level used, and reliable stabilization energies are only obtainable at high quantum-chemical ab initio levels. This is especially important for comparison of different structural types, such as H-bonded versus stacked structures. A reliable stabilization energy ($\Delta E$) follows from the complete basis set (CBS) limit of the CCSD(T) stabilization energy. Direct determination of this energy over an extended range is impractical, and the energy is thus approximated by the sum of the CBS limit of the MP2 stabilization energy ($\Delta E_{\text{MP2}}^{\text{CBS}}$) and a ($\Delta E_{\text{CCSD(T)}}^{\text{CBS}} - \Delta E_{\text{MP2}}^{\text{CBS}}$) correction term covering the difference between the MP2 and CCSD(T) stabilization energies (31). The latter term depends much less on the basis-set size and thus requires a rather small basis set:

$$\Delta E = \Delta E_{\text{CCSD(T)}}^{\text{CBS}} = \Delta E_{\text{MP2}}^{\text{CBS}} + (\Delta E_{\text{CCSD(T)}}^{\text{CBS}} - \Delta E_{\text{MP2}}^{\text{CBS}}).$$

Whereas we can already consider the Hartree-Fock (HF) interaction energy converged with respect to the one-electron basis set for relatively small basis sets, the MP2 correlation part of the interaction energy converges to its CBS limit unsatisfactorily slowly. Therefore, the HF and MP2 energies should be extrapolated separately. Several studies have demonstrated successful extrapolation schemes in the literature.
including the following equation by Helgaker and coworkers (38):

\[ E_{X}^{\text{HF}} = E_{\text{CBS}}^{\text{HF}} + A e^{-\alpha X} \quad \text{and} \quad E_{X}^{\text{corr}} = E_{\text{CBS}}^{\text{corr}} + B X^{-3}. \]

where \( E_{X} \) and \( E_{\text{CBS}} \) are energies for the basis set with the largest angular momentum \( X \) and for the CBS, respectively, and the \( \alpha \) parameter is fitted by the authors. The scheme uses the two-point form: It extrapolates two successive basis-set results. The two-point extrapolation form is preferable as inclusion of additional (lower-quality basis set) points results in extrapolations of lower quality to the fit, especially when using the smallest basis set (e.g., cc-pVDZ). The authors applied the basis-set superposition-error counterpoise correction and frozen-core approximations throughout this study. They also applied the CBS extrapolation to all calculated energies (dimer, monomers in both monomer- and dimer-centered basis sets, and monomers in vacuo); i.e., both deformation and basis-set superposition-error corrections were extrapolated.

So far we have considered only highly correlated wave-function theories. The use of popular DFTs cannot be recommended because they neglect the London dispersion energy part of the correlation (39–42). At present, the only possible approach is an empirical one in which an empirical dispersion energy augments the DFT energy. The approximate self-consistent charge density-functional tight-binding method, which is augmented by an empirical London dispersion energy, appears promising (43). The method yields good estimates of stabilization energies for stacked and H-bonded DNA base pairs and amino-acid pairs. The procedure is fast and can even be used in MD simulations.

### 3.1.4. Thermodynamic characteristics

MD/Q simulations with kinetic energy higher than any energy barrier scan the whole PES and localize all the energy minima. With sufficiently long simulations, the population of each structure can be found. Population is directly proportional to the change in free energy, and this approach thus extends the PES to the FES. Entropy always plays an important role, and certain types of energy minima are entropically favored. This is the case, for example, for stacked structures that are entropically favored over planar H-bonded structures. The MD/Q procedure is based on empirical potentials, but it is anharmonic in nature.

Another possibility for generating thermodynamic parameters is a standard statistical-thermodynamical treatment based on partition functions using the rigid-rotor harmonic-oscillator ideal-gas approximation. In this case, molecular parameters, such as rotational constants and vibrational frequencies, are obtained through a quantum-chemical treatment and are thus of higher quality, based on empirical potentials only. In addition to entropy, the enthalpy should be also determined. It is constructed as a sum of the interaction energy, the zero-point vibration energy, and temperature-dependent enthalpy terms.

### 3.1.5. Vibrational frequencies

Harmonic vibrational frequencies are computed by analytical or numerical calculation of second derivatives of the energy. The numerical approach is more time consuming, as well as less accurate, but it can be performed even for extended clusters described by large basis sets. The lowest level is formed by
correlated MP2 calculations using DZ + P basis sets. To match experimental frequencies, it is necessary to cover anharmonic effects, and the easiest way to achieve this is by introducing scaling factors. Satisfactory vibrational frequencies are mostly obtained when using a single-scaling factor based on experimental vibrational frequencies of the X-Y stretching coordinate of an isolated monomer. Use of multiple-scaling factors based on various X-Y coordinates should result in even closer agreement with experimental values.

### 3.2. Isolated Peptides

Isolated peptides exhibit features similar to those that characterize molecular clusters, and likewise two main structural motifs (H-bonding and stacking) dominate. The theoretical description of isolated peptides is therefore similar to that of molecular clusters, and only highly accurate quantum-chemical procedures yield satisfactory results in comparison with experiments.

#### 3.2.1. Strategy

The PES of simple peptides is considerably more complicated than that of DNA base pairs. Once again, stabilization of these structures is mainly a result of H-bonding and stacking, which involves delocalized electrons of phenyl groups, peptide bonds, and other units. The choice of empirical potential is difficult, and routinely used potentials such as AMBER or CHARMM are not always suitable. The aim is to select several dozen lowest-energy structures that can later be treated by reliable correlated ab initio optimization. This treatment is time consuming and therefore cannot be performed for all structures of a peptide. For example, in the case of Phe-Gly-Gly tripeptide, more than 1000 energy minima exist. The first screening method should be fast enough to allow MD simulations to be performed, and the use of the approximate self-consistent charge density-functional tight-binding method covering the dispersion energy is promising.

### 4. AMINO ACIDS AND PEPTIDES

#### 4.1. Amino Acids

There are three aromatic amino acids that provide the basic repertoire of chromophores for the study of peptides by electronic spectroscopy: phenylalanine (Phe), tyrosine (Tyr), and tryptophan (Trp). Their functional groups contain the respective chromophores of benzene, phenol, and indole. The spectra of these amino acids are affected by their molecular environment and therefore can report information on molecular conformation (44). Several authors have reported studies of derivatives of the three basic chromophores, such as tryptamine (45-48).

The conformers of several nonaromatic amino acids such as glycine (Gly), alanine (Ala), valine (Val), and proline (Pro) have been characterized using Fourier-transform microwave spectroscopy in a supersonic jet (49, 50). Typical stabilizing interactions derive from hydrogen bonds involving the acidic proton, the amino group, and parts of the functional group, such as π-electrons on an aromatic ring.
Fourier-transform IR spectroscopy of amino acids has been implemented recently by fast thermal heating of the amino-acid solids followed by fast-scan Fourier-transform IR prior to decomposition (51). At the high temperatures of the experiment, no conformer freezing is possible, but a large spectral range from near- to mid-IR is accessible so that the quality of vibrational calculations can be checked. Gly is the best-investigated amino acid with regard to the anharmonicity of its vibrations, and it can act as a benchmark when comparing the vibrational spectroscopy of amino acids with theory (52).

All amino acids investigated in the gas phase exhibit a neutral structure with COOH and NH terminal groups rather than the zwitterionic structure of the solids and the aqueous phase. Experiments to stabilize the zwitterionic form in the gas phase by sequentially adding water molecules have thus far been in vain (53, 54).

### 4.2. (Poly)amino Acids

(Poly)amino acids serve as model systems for the fundamental interactions that govern peptide folding with conformational motifs that are similar to folding motifs. As shown in Figure 3, we can categorize shapes based on the number of residues between hydrogen bonds.

One promising approach is the study of small segments of peptides, represented by very short sequences with capped end groups (55). We describe examples of these in the next section. In this section, we describe studies of small uncapped (poly)amino acids. When studying such sequences as a function of increasing length (such as X, X-Gly, and X-Gly-Gly), one finds that the propensity appears to be a reduction of the observed number of conformations with increasing length in spite of the larger number of possible shapes. For example, the six conformers of Trp reduce to four Trp-Gly and two Trp-Gly-Gly conformers (4). Presumably the increased length of the backbone allows the peptides to better lock in the most favorable conformations with the maximum number of stress-free hydrogen bonds and dispersion interactions. The most stable conformer of Trp-Gly-Gly, for example, is folded with an extended chain of hydrogen bonds, NH(indole)···O=C(R)-OH···O=C(R)-NH···O=C(γ-turn), and a favorable dispersion interaction between the side chain

![Figure 3](image_url)

Figure 3
Schematic classification of α-, β-, and γ-turn motifs.
Table 1 (Poly)amino acids studied by double-resonance spectroscopy

<table>
<thead>
<tr>
<th>Uncapped</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GW, WG, WGG</td>
<td>4, 18, 19, 116</td>
</tr>
<tr>
<td>WAGGDASE</td>
<td>19</td>
</tr>
<tr>
<td>SW</td>
<td>116</td>
</tr>
<tr>
<td>PW</td>
<td>116</td>
</tr>
<tr>
<td>PG, PGG</td>
<td>25, 117</td>
</tr>
<tr>
<td>YA, AY</td>
<td>117</td>
</tr>
<tr>
<td>LF, PP</td>
<td>56</td>
</tr>
<tr>
<td>Cyclo PS</td>
<td>57</td>
</tr>
<tr>
<td>Gramicidin A-D, S</td>
<td>58</td>
</tr>
<tr>
<td>WGGGY</td>
<td>58</td>
</tr>
<tr>
<td>FDASV</td>
<td>118</td>
</tr>
<tr>
<td>Cyclo GY</td>
<td>69</td>
</tr>
</tbody>
</table>

α-, β-, γ-turns: certain turns, involving hydrogen bonding between specific amino-acid intervals and leading to typical folding motives and the aromatic ring system (18). Table 1 lists (poly)amino acids that have been studied by double-resonance spectroscopy.

The importance of dispersion poses an extra challenge for the theoretical treatment of these structures, as discussed in Section 3 (see Figure 2). Geometrical restraints may hinder stress-free hydrogen bonds so that dispersion stabilization can compete favorably. Several tripeptides exhibit γ-turn-type structures, and α-turn-type structures have been observed in pentapeptides. Structures of cyclic and chiral dipeptides have been reported as well (56, 57). There are also reports of results on peptides of increasing size, including the nine-residue delta-sleep-inducing peptide (19) and 15-residue gramicidins (58).

Figure 4 shows IR-UV spectra of three peptides that represent different secondary-structural folding patterns, a random-coil Gly sequence, cyclic gramicidin S that forms internal bonds similar to a β-sheet structure, and gramicidin A with a β-helix structure. These experiments give hope that the first steps of peptide folding can be studied in the gas phase in microscopic detail. Entropic effects, however, are probably less important at the ultralow temperatures of the jet experiments so that the jet temperature has to be varied and microhydrated peptides have to be studied to elucidate hydrophobic/hydrophilic (entropic) effects during peptide folding. No zwitterionic peptides have been observed in the gas phase.

4.3. Protected End Groups

A number of groups have studied amino acids and peptides with so-called protected end groups (acetylated, methylated, and cyclic). A recent review by Chin et al. (55) provides a detailed account of these studies, so we merely summarize them here. Capped peptides can often be transferred to the gas phase by simple thermal heating. They are cutouts of structural motifs in large peptides in which the end groups do not significantly contribute to global folding. γ-turns have been characterized...
**Figure 4**

Infrared-ultraviolet spectral hole burning (IR-UV) spectra of three peptides that represent different secondary-structural folding patterns: (a) random-coil glycine sequence, (b) gramicidin A exhibiting a β-helix structure, and (c) cyclic gramicidin S forming internal bonds similar to a β-sheet structure.

by SHB (18, 19, 59–62). These turns and loops do not occur periodically but at well-defined locations in proteins, and they determine their folding just as α-helices and β-sheets do.

By the dimerization of stretched conformers of protected amino acids and peptides, researchers have formed β-sheet model systems in the gas phase with intermolecular C=O···HN hydrogen bonds (20, 61, 63, 64). These β-sheets can be stabilized by multipoint connectors such as pyrazole. These studies may support the development of optimized drugs against diseases such as Alzheimer’s (64).

**4.4. Conformer-Dependent Dynamics**

The transition-moment orientation and S1 lifetime of Phe strongly depend on the conformation of the alanyl side chain (21). The most stable conformer with a COOH → NH2 → π H-bond sequence has a lifetime of 20 ns compared with 80–120 ns for the other conformers, revealing an efficient nonradiative-decay pathway (21). The transition-moment orientation of tryptamine drastically depends on vibronic excitation of the respective conformer, indicating a vibration- and
conformation-dependent coupling of the initially excited 1Lb(S1−) state to other nearby states, possibly the 1La(S2−) state (65).

Sobolewski and coworkers (66) have proposed an internal conversion (IC) mechanism via a charge-transfer intermediate state that may govern excited-state dynamics in many H-bonded biomolecules, including peptides. We discuss this mechanism further in Sections 5.1 and 5.2.

Lee et al. (67, 68) measured the ionization energies of Phe using tunable two-color resonant two-photon ionization and found them to be strongly conformation dependent. The most stable conformers change their structure drastically on ionization because the attractive interaction between the amino hydrogen atom and the phenyl ring turns into a repulsive one (68), which leads to the higher energy of the cation. Similarly, the most stable conformer of cyclic Trp-Gly exhibits a peptide ring to indole ring interaction, leading to a repulsion between the two rings in the cation and a higher ionization energy than that of 3-methylindole (69).

Peptide radical cations can exhibit photo-activated positive-charge migration from the initially electronically excited phenyl, tyrosyl, and indole cation to the N terminus, as identified by a positively charged N-terminal fragment (70). The efficiency of this charge-transfer process depends not only on the number and type of residues between the excited chromophore and the N terminus, but also on the amino-acid or peptide conformation (67, 68, 71, 72).

5. DNA

5.1. Single Bases

Brady et al. (16) reported the first attempt to record the electronic spectroscopy of DNA bases in the gas phase with the broad spectra for the pyrimidine bases uracil and thymine. Nir et al. (73) demonstrated the first successful experiment to report a well-resolved DNA base spectrum, by recording the discrete electronic spectrum of laser-desorbed, jet-cooled guanine. This was followed by a second purine, adenine (74), and finally by the pyrimidine cytosine, as well as derivatives of these bases (75). Nir et al. (76) also reported the electronic spectroscopy of the nucleoside guanosine. Two major issues emerged for these molecules: The spectra are complicated by tautomerism, and their photochemistry appears dominated by the properties of the excited electronic state.

5.1.1. Tautomers. Nucleobases can adopt a number of different tautomeric forms. Tautomerism has been implicated in theories describing mutagenic reactions. The four lowest-energy tautomers of the purine bases are the keto and enol forms, which can each appear in the N7H and the N9H form. Figure 5 shows an example of the four lowest-energy tautomers of guanine, all of which have been found in the gas phase (77, 78). Enol and keto tautomers are easily distinguished by IR-UV SHB, thanks to the characteristic enol-OH stretch at approximately 3600 cm⁻¹. However, because the N7H and N9H frequencies are quite similar, some uncertainty remains regarding the final assignment of these forms. Comparison with the origins of tautomERICALLY
WC: Watson-Crick

Base pairing: recognition in DNA relies on the preferred base-pair formation of cytosine-guanine and adenine-thymine by hydrogen bonding.

Figure 5

The four lowest-energy tautomers of guanine.

Figure 5

The four lowest-energy tautomers of guanine.

The existence of keto and enol tautomers of neutral cytosine in the gas phase had already been identified by anion spectroscopy (80). In the gas phase, the nucleobases are found in both the keto and enol forms. All nine substituted derivatives of guanine, including the guanosine nucleosides, are observed exclusively in the enol form in the gas phase when using REMPI or LIF for analysis (81). This is intriguing because in solution the nucleosides appear exclusively in the keto form, required for the Watson-Crick (WC) base pairing in DNA. Failure to observe these keto tautomers may be a result of a short excited-state lifetime, as we discuss below. Regarding the other nucleobases, N9-H adenine (82) and diketo-thymine (83) seem to be the dominant tautomers in the gas phase.

5.1.2. Photochemistry. The nucleobases involved in replication have a short excited-state lifetime (84). This may provide a mechanism to protect the building blocks of life against photochemical damage, by providing a pathway for rapid IC to
the ground state (85). In this way, the molecule can dispose of excitation energy to its environment in the form of heat, rather than being trapped in a reactive excited state, with even possible intersystem crossing to a triplet state. The proposed mechanism involves a so-called doorway state with curve crossings or conical intersections, connecting the S1 excited state and the S0 ground state. The nature of this doorway state has been the subject of a number of investigations. Broo (85) has proposed that it might consist of a state with nπ* character and that the relative energies of the nπ* and ππ* transitions are critical in determining the photochemistry. More recent models (86, 87) include mixed nπ*/ππ* states owing to puckering of the six-membered ring of adenine at the C2-H group. Consistent with this mechanism, 2-aminopurine has a large fluorescence quantum yield, as opposed to its isomer adenine (88–90). A weak absorption can also be observed for adenine close to that of the dominant spectrum, which can be ascribed to the nπ* transition, consistent with this model (74, 82). Sobolewski and coworkers (66, 91) have proposed a doorway state of πσ* character, associated with motion along an N-H coordinate; for adenine this would be the N9-H coordinate. Consistent with this mechanism, Hüning et al. (92) directly observed N-H photodissociation in adenine at 243-nm photolysis energy via the resonant ionization of product H atoms arising from N9-H but also to some extent from the NH2 group. Zierhut et al. (93) obtained similar results at 239-nm and 266-nm photolysis energy. Conversely, Kang et al. (94) measured the excited-state lifetimes in femtosecond 267-nm pump-probe experiments and found them to be similar for all adenine derivatives, including 9-methyl adenine. The onset of the πσ* state therefore is probably at approximately 266-nm or even shorter wavelengths, and a contribution to fast IC at 277–273 nm in the region of the sharp vibronic resonances of adenine is improbable.

The IC pathways of the other nucleobases are less well investigated. They seem to involve C=C twisting in cytosine (95), with a low barrier for keto-cytosine and a considerable energy gap to the S0 state even at extended twisting for enol-cytosine (96). For 9-methyl guanine, the IC pathway seems to involve a strongly bent amino group in position two (97).

5.2. Base Pairs

Cluster formation in a supersonic expansion provides the opportunity to study interactions between individual molecules. Purines and pyrimidines have been studied as base-pair mimics (98). The first reported REMPI spectrum of an actual base-pair cluster involved guanine-cytosine (GC) (99). As of this writing, the spectra have been reported for over 27 base-pair combinations, with REMPI, UV-UV, and IR-UV double resonance for most of these (96) (see Table 2).

Anharmonic vibrational frequencies were recently determined for the GC complex and compared with gas-phase IR-UV double-resonance spectral data. Harmonic frequencies were obtained at the RI-MP2/cc-pVDZ and RI-MP2/TZVPP (triple-zeta valence double-polarization basis-set) levels, and anharmonic frequencies were obtained by the vibrational self-consistent field method based on the improved semiempirical parameterized method 3 results. Comparison of the data with
Table 2 Origins of selected nucleobases and their clusters in cm<sup>-1</sup>

<table>
<thead>
<tr>
<th>Bases</th>
<th>Origins (cm&lt;sup&gt;-1&lt;/sup&gt;) of different structures</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>32,864; 33,269; 33,928; 34,773</td>
<td>73, 77, 121, 122</td>
</tr>
<tr>
<td>7mG</td>
<td>32,579; 33,080</td>
<td>121</td>
</tr>
<tr>
<td>9mG</td>
<td>34,630</td>
<td>121</td>
</tr>
<tr>
<td>A</td>
<td>36,105</td>
<td>74, 89</td>
</tr>
<tr>
<td>C</td>
<td>31,826</td>
<td>75, 79</td>
</tr>
<tr>
<td>1mC</td>
<td>31,908</td>
<td>96</td>
</tr>
<tr>
<td>5mC</td>
<td>31,269</td>
<td>96</td>
</tr>
<tr>
<td>Guanosine</td>
<td>34,492</td>
<td>76, 81</td>
</tr>
<tr>
<td>Adenosine</td>
<td>35,550</td>
<td>123</td>
</tr>
<tr>
<td>2-aminopurine</td>
<td>32,368</td>
<td>89</td>
</tr>
<tr>
<td>G-H&lt;sub&gt;2&lt;/sub&gt;O</td>
<td>33,049; 33,218; 33,301</td>
<td>108</td>
</tr>
<tr>
<td>G-G</td>
<td>33,103; 33,282</td>
<td>125</td>
</tr>
<tr>
<td>C-C</td>
<td>33,483</td>
<td>124</td>
</tr>
<tr>
<td>G-C</td>
<td>32,810; 33,314</td>
<td>119, 120</td>
</tr>
<tr>
<td>A-T</td>
<td>35,064</td>
<td>104, 114</td>
</tr>
<tr>
<td>GG(H&lt;sub&gt;2&lt;/sub&gt;O)</td>
<td>32,926</td>
<td>126</td>
</tr>
<tr>
<td>A-A</td>
<td>35,040</td>
<td>127</td>
</tr>
<tr>
<td>C-1mC</td>
<td>33,419</td>
<td>124</td>
</tr>
<tr>
<td>C-5mC</td>
<td>32,500; 32,691; 32,916</td>
<td>124</td>
</tr>
<tr>
<td>5mC-5mC</td>
<td>32,493; 32,691; 32,872</td>
<td>124</td>
</tr>
</tbody>
</table>

Multiple values are for multiple isomers.

The experimental results indicate that the average absolute-percentage deviations for the method are 2.6% for harmonic RI-MP2/cc-pVDZ, 2.5% for harmonic RI-MP2/TZVPP, and 2.3% for adopted PM3 CC-VSCF (parameterized method 3 vibrational self-consistent field). The use of an empirical scaling factor for the ab initio harmonic calculations improves the stretching frequencies but decreases the accuracy of the other mode frequencies. The mid-IR region can also provide additional information about sugar conformations in the study of nucleoside clusters (100). The REMPI spectra show low-frequency H-bond vibrations, as expected for the six H-bond modes. Those mode frequencies are not diagnostic for structure because their order of magnitude is similar in most cases. Stacked structures are not observed, with the exception of methylated adenine dimers (101). In this case, the derivatization blocks H-bonding sites of the lowest-energy H-bonded structures.

Possibly the most intriguing finding is that all observed combinations of base pairs so far exhibit sharp REMPI spectra, with the notable exception of the three cases assigned as WC structures (see Figure 6). In fact, the two non-WC GC structures feature sharp REMPI spectra, whereas the WC structures of the same, albeit derivatized, base pair are broad. A possible explanation involves a similar model as that for the short-lived monomers, namely a coupling with a doorway state that in turn couples with the electronic ground state via conical intersections. Depending on the
energetics of the excited states, the transition from the S1 state to the doorway state may be barrierless, leading to a diffuse spectrum, or may feature a barrier, resulting in a discrete spectrum. The nature of the doorway states may be different from that of the monomers. Sobolewski and colleagues (102) predict a charge-transfer state that indeed provides a barrierless transition from S1 and a conical intersection, with S0 only for the WC structure. One intriguing implication is that the structure that rapidly reconverts to the ground state is more photochemically stable. This may have been especially relevant under the conditions of prebiotic chemistry with abundant deeper UV irradiation than today, before the formation of the atmospheric ozone layer. Importantly, in solution this particular decay mechanism of DNA competes with charge transfer along the stacked structure.

Short-lived excited states reduce multiphoton ionization efficiency. Femtosecond excited states are dark to REMPI with nanosecond laser pulses. Therefore, there may

Figure 6
Data for isolated guanine-cytosine (GC) base-pair structures. (a) Results for the Watson-Crick structure. (b,c) The second and third lowest-energy structures, respectively. The second column shows the infrared-ultraviolet (IR-UV) double-resonance data compared with the ab initio calculations. Lines with an asterisk refer to a mode localized on the guanine moiety. Broad red-shifted peaks are the result of hydrogen bonding. The third column shows the UV-excitation spectra, measured by resonant two-photon ionization (R2PI). The fourth column shows potential energy diagrams from Reference 102.
Microhydration: solvation by a small number of water molecules to study solvation details

be isomers in the molecular beam that go unobserved by virtue of the mismatch of timescales of the detection method and excited-state lifetime. Structures that so far have not been observed, in spite of predicted low energy, include symmetric GG and AA dimers, keto tautomers of H9 substituted guanines, and, as discussed below, various H2O clusters.

The relative stability of clusters is one of the most valuable parameters for comparison with theory; however, it is also one of the most difficult ones to obtain experimentally. Jet cooling is a nonequilibrium process, and relative abundances of isomers in a molecular beam are thus not reliable observations for determining thermodynamic quantities. So far only field ionization measurements have provided some experimental data on the relative stability of the structures (103).

5.3. Complexes with Water

One appealing feature of molecular-beam spectroscopy is the ability to study mass-selected clusters with water. This so-called microhydration makes it possible to observe the role of the solvent by sequentially adding one water molecule at a time. Kim et al. (104) employed a specially designed thermal-evaporation source to measure the ionization potentials of adenine and thymine clusters as a function of the number of water molecules. Water also plays an important role in the structure and function of DNA and RNA. A number of theoretical studies aimed at describing the effects of water molecules on the purine nucleobases (105–107). Because of the relative difficulty of bringing the nucleobases into the gas phase and clustering them with water, few experimental studies examine the effects of microscopic water on the structure of the bases (106, 108). Kim and colleagues (109, 110) concluded that electronically excited clusters of the adenine monomer with water dissociate within 200 fs as a result of vibronic coupling with a repulsive state of n-π* character. He et al. (111) examined the effects of water molecules on the S1 lifetimes of the pyrimidine bases and concluded that the presence of water molecules could increase the IC rate to the ground state through an intermediate dark state. Chin et al. (106) observed just one structure, an enol-amino tautomer, for 9-methylguanine monohydrate under jet-cooled conditions, although calculations predict a keto structure to be close in energy, just 0.45 kcal mol⁻¹ higher than the enol tautomer. Crews et al. (108) found three structures for guanine monohydrate, including both enol and keto forms.

Theoretical calculations show that for the adenine-thymine base pair, the presence of two water molecules suffices to stabilize a stacked structure more than a H-bonded structure (112, 113). The enhanced stabilization results from the water molecules surrounding the base pair and creating a bridge between the bases. Calculations predict that for GG base pairs, approximately 20% of the structures stack on addition of two water molecules (112). Additional water further stabilizes stacking structures, and on the addition of a fourth water molecule, 90% of the structures are predicted to be stacked. Sivanesan et al. (114) predict that five to six water molecules preferentially stabilize stacked GC pairs. Microhydration also can stabilize specific tautomers (107). IR-UV resonance spectra by Abo-Riziq et al. (115) on microhydrated GG pairs suggest that a single H2O molecule suffices to stabilize one specific base-pair structure,
relative to one that in the absence of solvent is close in energy. Two water molecules do not suffice to lead to observable populations of a stacked GG structure.

6. OUTLOOK
In the past five years, researchers have made accelerated progress in the gas-phase spectroscopy of biomolecular building blocks. This progress has been driven in part by the confluence of improved capabilities in studying molecules of increasing size in both experiment and theory. In this lucky circumstance, gas-phase experimental data serve as benchmarks for computational algorithms and force fields, whereas computational results help interpret and guide experiments. As a result, an increasing number of research groups has entered this field, and significant further progress may be expected in years to come. Challenges include the following: (a) increasing the size of the molecules studied, (b) expanding the studies of larger water clusters to serve as a bridge with the solution state, and (c) studies of the dynamics of isolated systems, such as charge transfer, proton transfer, and isomerization. Fast excited-state dynamics may necessitate studies at subpicosecond timescales.

SUMMARY POINTS
1. The study of isolated biomolecules provides insights into intrinsic properties that are otherwise masked by elements of the biological environment, such as solvent and macromolecular structure. Gas-phase experiments, together with quantum calculations, reveal details of structure, interactions, and dynamics at the molecular level.
2. Theoretical treatment of biomolecular building blocks of increasing size requires the development of new computational strategies. Among the special challenges are large conformational landscapes and the role of dispersive forces.
3. Hydrogen bonding and dispersive forces are the two major motifs stabilizing biomolecular structure. An example is the combination of base pairing and stacking in DNA.
4. Exploration of the conformational landscape of (poly)amino acids shows elements of folding motives in peptides.
5. Clusters of DNA bases reveal the unique hydrogen bonding that plays a role in base pairing.
7. The study of clusters with water provides insight into microhydration, revealing details of solvation.
ACKNOWLEDGMENTS

This material is based on work supported by the National Science Foundation under grant CHE-0244341 and by the Donors of the American Chemical Society Petroleum Research Fund. This work was supported by grants 2003/05/0009, A400550510, and LC512 from the Grant Agency of the Czech Republic, Grant Agency of the Academy of Sciences of the Czech Republic, and MSMT of the Czech Republic.

LITERATURE CITED


25. Presents a computational strategy to deal with dispersive forces.


44. Explores computational landscapes.

58. Presents a gas-phase study of 15 amino-acid peptides.
59. Explores gas-phase structure and folding motives.


Contents

Frontispiece

C. Bradley Moore ................................................................. xvi

A Spectroscopist's View of Energy States, Energy Transfers, and Chemical Reactions
C. Bradley Moore .................................................................. 1

Stochastic Simulation of Chemical Kinetics
Daniel T. Gillespie .................................................................. 35

Protein-Folding Dynamics: Overview of Molecular Simulation Techniques
Harold A. Scheraga, Mey Khalili, and Adam Liwo ..................... 57

Density-Functional Theory for Complex Fluids
Jianzhong Wu and Zhidong Li .................................................. 85

Phosphorylation Energy Hypothesis: Open Chemical Systems and Their Biological Functions
Hong Qian .............................................................................. 113

Theoretical Studies of Photoinduced Electron Transfer in Dye-Sensitized TiO2
Walter R. Duncan and Oleg V. Prezhdo .................................... 143

Nanoscale Fracture Mechanics
Steven L. Mielke, Ted Belytschko, and George C. Schatz ............ 185

Modeling Self-Assembly and Phase Behavior in Complex Mixtures
Anna C. Balazs ....................................................................... 211

Theory of Structural Glasses and Supercooled Liquids
Vassily Lubchenko and Peter G. Wolynes ................................. 235

Localized Surface Plasmon Resonance Spectroscopy and Sensing
Katherine A. Willets and Richard P. Van Duyne ...................... 267

Copper and the Prion Protein: Methods, Structures, Function, and Disease
Glenn L. Millhauser .................................................................. 299
Aging of Organic Aerosol: Bridging the Gap Between Laboratory and Field Studies
Yinon Rudich, Neil M. Donahue, and Thomas F. Mentel .................................................. 321

Molecular Motion at Soft and Hard Interfaces: From Phospholipid Bilayers to Polymers and Lubricants
Sung Chul Bae and Steve Granick .......................................................... 353

Molecular Architectonics on Metal Surfaces
Johannes V. Barth .................................................................................. 375

Highly Fluorescent Noble-Metal Quantum Dots
Jie Zheng, Philip R. Nicovich, and Robert M. Dickson .......................... 409

State-to-State Dynamics of Elementary Bimolecular Reactions
Xueming Yang .................................................................................. 433

Femtosecond Stimulated Raman Spectroscopy
Philipp Kukura, David W. McCamant, and Richard A. Mathies ............. 461

Single-Molecule Probing of Adsorption and Diffusion on Silica Surfaces
Mary J. Wirth and Michael A. Legg ..................................................... 489

Intermolecular Interactions in Biomolecular Systems Examined by Mass Spectrometry
Thomas Wyttenbach and Michael T. Bowers ........................................ 511

Measurement of Single-Molecule Conductance
Fang Chen, Joshua Hihath, Zhifeng Huang, Xiulan Li, and N.J. Tao ........ 535

Structure and Dynamics of Conjugated Polymers in Liquid Crystalline Solvents
P.F. Barbara, W.-S. Chang, S. Link, G.D. Scholes, and Arun Yethiraj ....... 565

Gas-Phase Spectroscopy of Biomolecular Building Blocks
Mattanjah S. de Vries and Pavel Hobza ................................................ 585

Isomerization Through Conical Intersections
Benjamin G. Levine and Todd J. Martinez .......................................... 613

Spectral and Dynamical Properties of Multiexcitons in Semiconductor Nanocrystals
Victor I. Klimov ................................................................................ 635

Molecular Motors: A Theorist’s Perspective
Anatoly B. Kolomeisky and Michael E. Fisher .................................... 675

Bending Mechanics and Molecular Organization in Biological Membranes
Jay T. Groves .................................................................................. 697

Exciton Photophysics of Carbon Nanotubes
Mildred S. Dresselhaus, Gene Dresselhaus, Riichiro Saito, and Ado Jorio ...... 719