

# The Chemist's Concept of Molecular Structure

*N. Sukumar*

*Department of Chemistry & Biological Chemistry and*

*Center for Biotechnology & Interdisciplinary Studies*

*Rensselaer Polytechnic Institute*

*Troy, NY*

*Email: [nagams@rpi.edu](mailto:nagams@rpi.edu)*

*Phone: (518)276-4235*

*Fax: (518)276-4233*

## **ABSTRACT:**

The concept of molecular structure is fundamental to the practice and understanding of chemistry, but the meaning of this term has evolved and is still evolving. The Born-Oppenheimer separation of electronic and nuclear motions lies at the heart of most modern quantum chemical models of molecular structure. While this separation introduces a great computational and practical simplification, it is neither essential to the conceptual formulation of molecular structure nor universally valid. Going beyond the Born-Oppenheimer approximation introduces new paradigms, bringing fresh insight into the chemistry of fluxional molecules, proteins, superconductors and macroscopic dielectrics, thus opening up new avenues for exploration. But it requires that our ideas of molecular structure need to evolve beyond simple ball-and-stick-type models.

## Why should molecules have structure?

*“The idea that molecules are microscopic, material bodies with more or less well-defined shapes has been fundamental to the development of our understanding of the physicochemical properties of matter, and it is now so familiar and deeply ingrained in our thinking that it is usually taken for granted - **it is the central dogma of chemistry.**”*

— R. G. Woolley (Woolley 1980)

The notion that a molecule has structure is fundamental to much of chemistry as practiced today. But what do we really mean by this term? The concept of molecular structure has evolved over the course of the past century and a half (Ackeroyd 2005) and is still evolving today. There are several ways that we can envision molecular structure, some being more general and others fairly concrete, but rather restrictive. As an example of the latter, we can think of molecular structure in terms of familiar ball-and-stick-type molecular models. Such models are simple to visualize and are intuitively appealing. But by confining our conception to such models, we risk imposing a classical, mechanical, vision upon an intrinsically microscopic quantum world.<sup>1</sup> Alternatively, we can define structure as that property of matter by virtue of which it occupies space in the real world. In other words, structure is what prevents a molecule, or any matter, from collapsing onto

---

<sup>1</sup> Brown (Brown 2003a; Brown 2003b) has argued that molecular structure models should be seen as metaphors that reflect our conceptualization of chemistry and the manifestation of subconscious forms of understanding in the process of scientific reasoning.

itself. From a statistical perspective, we can define structure as that which distinguishes an object from a heap of its parts, in this case, a molecule from a collection of its constituent atoms.

The statistical definition given above generalizes the concept of molecular structure to situations where the relative spatial locations of the constituent atoms may not be fully known, or even well-defined. Indeed, most modern molecular structure determinations are indirect, utilizing a transformation from momentum space, or from frequency domain, to position space. Structure is quantified by the inter-particle distribution function.<sup>2</sup> In this view, molecular structure is tantamount to correlations between geometric parameters describing the molecule. Thus an ideal gas of atoms has minimal structure, a hydrogen-bonded liquid is more structured and a crystal or molecular solid even more so. The familiar ball-and-stick molecular models are thus seen to be the rigid limit of a hierarchy of structures (Smith 1981; Eberhart 2002). It is worth bearing this in mind as

---

<sup>2</sup> Structure can be quantified by structure factors, inter-particle distribution functions or pair correlation functions. Thus the two-particle density is given by:

$$\gamma(r, r') = \int dr_3 \int dr_4 \dots \int dr_N \psi^*(r, r', r_3, r_4, \dots r_N) \psi(r_3, r_4, \dots r_N)$$

where  $\psi(r, r', r_3, r_4, \dots r_N)$  is the quantum wave function as a function of the coordinates of all particles in the system,  $\psi^*$  its complex conjugate and the integrals run over the coordinates of all particles but two. The two-particle density can further be integrated over angular coordinates to give the radial distribution function, a function of a single distance. See Fig.2 for illustrative examples.

we prepare to move beyond the Born-Oppenheimer approximation (BOA) to explore Jahn-Teller systems, transition states, fluxional molecules and proteins.

The BOA (Born and Oppenheimer 1927; Born 1951; Born and Huang 1954) is central to the modern chemists' conception of molecular structure and forms a powerful paradigm underlying many chemical theories. There has been much discussion and debate over the years (Woolley 1976; Lathouwers, VanLeuven et al. 1977; Woolley and Sutcliffe 1977; Lathouwers and VanLeuven 1978; Woolley 1978; Woolley 1978; Bader, Nguyen-Dang et al. 1979; Trindle 1980; Woolley 1980; Bader, Nguyen-Dang et al. 1981; Bader 1990; Sutcliffe 1992) as to whether the concept of molecular structure survives beyond BOA. Woolley (Woolley 1976; Woolley and Sutcliffe 1977; Woolley 1978; Woolley 1978; Woolley 1980), Woolley and Sutcliffe (1977) has argued that the notion that a single molecule has a shape – *i.e.*, a nearly fixed (relative) arrangement of nuclei in space – is a classical idea imposed upon the quantum mechanical picture of matter. Obviously our ball-and-stick-type models do not survive when we move into realms where BOA breaks down, as indeed it does for Jahn-Teller molecules and other instances of curve-crossing systems. However, although very useful and powerful, the BOA is neither necessary nor sufficient for formulating a concept of molecular structure.

## **Molecular structure in the Born-Oppenheimer approximation:**

*“Approximations are the most precious fruits of theory.”*

— Valentin Ostrovsky (Ostrovsky 2005)

The original BO separation (Born and Oppenheimer 1927) of electronic and nuclear motions in molecules was formulated as a perturbation expansion in the ratio of the electronic to nuclear masses, to show that molecular states can be approximately represented as products of electronic and nuclear functions. The electronic structure problem then involves solving for the eigenfunctions of an electronic Hamiltonian, while the nuclear function satisfies an equation of motion, with the eigenvalues of the electronic Hamiltonian forming an effective potential energy surface upon which the nuclei may be envisioned to move. The distinct concepts of electronic structure and molecular structure are thus intimately related. This is, of course, not accidental: as Hohenberg and Kohn (Hohenberg and Kohn 1964) showed, there exists a unique mapping between the potential  $v(\mathbf{r})$  due to the nuclei and the distribution of electron density  $\rho(\mathbf{r})$ . Since  $\rho(\mathbf{r})$  determines the number of electrons

$$N = \int \rho(\mathbf{r}) \, d\mathbf{r} \quad (1),$$

it follows that  $\rho(\mathbf{r})$  also uniquely determines the ground state wave function  $\psi$ , the ground state electronic energy and the molecular structure.

The original perturbative BO treatment has given rise to considerable confusion and has since been re-examined and extended by several workers, starting with Born himself (Born 1951; Born and Huang 1954). It is now recognized that a necessary, though not sufficient, condition for the validity of BOA in a certain region of nuclear configuration space is that the electronic wave function be non-degenerate at all points of interest in that region. Furthermore, the nature of the intramolecular forces is of crucial importance in determining the qualitative features of molecular systems and in effecting the BO

separation: there is no separation of electronic and nuclear motions for a hypothetical system of particles with the same masses as in a molecule, but with all forces attractive, or for an analogous cluster of (heavy) stars with (light) planets.

In an interesting analysis of the BO separation, Essén (Essén 1977; Essén 1983) pointed out that, according to the virial theorem:

$$2\langle T \rangle = -\langle V \rangle \quad (2)$$

the potential energy ( $V$ ) is the dominant term in the total energy ( $T + V$ ); a low energy state can be thus attained only if  $V$  is strongly negative. This requires that the particles are arranged in essentially neutral atomic subsystems. The intramolecular motions can then be approximately decomposed into three nearly independent motions: (1) translational motion of the molecule as a whole, (2) collective motions of the neutral atomic subsystems, and (3) internal motions within each atom, which constitutes the major part of the potential energy of internal motion in the molecule. The separability of internal motions in molecules arises from the fact that within each strongly interacting atomic subsystem, the internal momenta of the electrons and nuclei are of comparable magnitude ( $mv_i \sim MV_i$  where  $m$  and  $M$  are electronic and typical nuclear masses,  $v_i$  and  $V_i$  are typical internal velocities of electrons and nuclei, respectively). Due to their small relative mass, the electrons contribute most of the kinetic energy of individual motions within the atomic subsystems ( $\frac{1}{2} mv_i^2 \gg \frac{1}{2} MV_i^2$ ), while the kinetic energy of atomic collective motions is largely determined by the nuclei. Upon separating collective and internal motions in a molecule, these can be approximately identified with relative

nuclear motions and electronic motions relative to fixed nuclei. This is how the mass ratio enters into the BO scheme.

### **Molecular structure in the Quantum Theory of Atoms in Molecules:**

*“Will you reflect for a moment on some of the things that I have been saying? I described a bond, a normal simple chemical bond; and I gave many details of its character (and could have given many more). Sometimes it seems to me that a bond between two atoms has become so real, so tangible, so friendly that I can almost see it. And then I awake with a little shock; for a chemical bond is not a real thing: it does not exist: no one has ever seen it, no one ever can. It is a figment of our own imagination.”*

— C. A. Coulson (Coulson 1951; Coulson 1955)

Virial partitioning of molecular systems into roughly neutral subsystems, as discussed in the previous section, forms the basis of the Quantum Theory of Atoms in Molecules (QTAIM) (Bader 1990; Matta and Boyd 2007). This theory provides a rigorous physics-based recipe for partitioning a molecule into atomic subsystems. In this formulation, the nuclei function as attractors of the electron density field  $\rho(\mathbf{r})$ , the atom being defined as the union of an attractor and its basin of attraction. Each atom thus contains one and only one nucleus, with the gradient paths of the electron density ( $\nabla\rho$ ) being employed to define the bonds between atoms as well as the interatomic boundaries: the bond path between any two atoms is defined as the unique gradient path  $\nabla\rho$  connecting the

respective nuclei, while the atomic surface is defined as the locus of points satisfying the zero-flux criterion:

$$\nabla\rho\cdot\hat{\mathbf{n}} = 0 \quad (3)$$

where  $\hat{\mathbf{n}}$  is the normal to the surface (Fig. 1). The basis for this partitioning scheme is that the zero-flux criterion ensures that each atomic subsystem satisfies the virial theorem, thereby giving rise to spatial additivity of the action

$$\mathcal{W} = \int \mathcal{L}(t) dt \quad (4)$$

(where  $\mathcal{L}$  is the Lagrangian) and of its variation, in accordance with Schwinger's principle of stationary action (Schwinger 1951). Schwinger's principle extends the formulation of quantum mechanics to open quantum subsystems, such as atoms in molecules. The concept of chemical bonds is recovered through bond paths in Bader's quantum theory: the topology of the bond paths completely specifies the molecular graph (Bader, Srebrenik et al. 1978; Bader, Anderson et al. 1979; Bader 1980; Bader, Nguyen-Dang et al. 1981; Bader 1990; Bader 1991; Bader 1995; Bader 1998; Bader 1998), also referred to as the 2-D structure. Fig.1 shows the correspondence between molecular electron densities, gradient vector ( $\nabla\rho$ ) maps and conventional conceptions of molecular structure for the case of benzene.

The chemical utility of Bader's QTAIM results from the approximate transferability of atomic and functional group properties (Bader, Popelier et al. 1994; Sukumar and Breneman 2006), defined through the zero-flux criterion (3), within similar molecules.

This feature enables a systematic study of chemical properties and is really what makes chemistry a science rather than a mere catalog of properties of disparate molecules. The philosophical utility of the QTAIM decomposition has been realized by several authors (Reiher 2003; Shahbazian and Zahedi 2006).

## **Molecular structure beyond BOA:**

*“Actually, there is no salt in the sea.”*

— Joseph E. Early (Early 2005)

The 2-D structure of a molecule, as represented by its molecular graph, is preserved over the course of normal molecular vibrations and rotations. When the molecular graph changes, the molecule is said to undergo dissociation, association or rearrangement to a different isomer. It is important to remember that 2-D structure, although derivable from QTAIM, generally forms the input to an electronic structure computation within BOA. There have been some efforts directed to first principles computation of molecular structure beyond BOA. I. L. Thomas (Thomas 1969; Thomas 1970; Thomas 1970; Thomas and Joy 1970; Thomas 1971; Thomas 1972) performed a pioneering series of electronic and protonic structure computations at the Hartree-Fock level, for a series of second-row hydrides, putting electrons into electronic orbitals and protons into protonic orbitals, separately antisymmetrizing the electronic and protonic wave functions and iterating the Fock equations to self-consistency. Equilibrium bond lengths and bond angles emerge in this picture as peaks in the protonic radial distribution function. The results were remarkable: the structure of ammonia, for instance, turned out to be a disk

centered at the nitrogen atom, apparently very different from the expected pyramidal structure!

While the pyramidal structure is indeed the equilibrium ground state structure of ammonia obtained from solution of the electronic Schrödinger equation within BOA, it is **not** an eigenfunction of the total Hamiltonian of the isolated molecule:

$$\mathbf{H} = \mathbf{H}_e + \mathbf{T}_n \quad (5)$$

(where  $\mathbf{H}_e$  is the electronic Hamiltonian and  $\mathbf{T}_n$  is the nuclear kinetic energy operator).

The ground state solution  $\Psi_+$  of the total molecular Schrödinger equation beyond BOA is a symmetric linear combination of two pyramidal ammonia wave functions ( $\Psi_L$  and  $\Psi_R$ ) that are related by a reflection operation (Fig. 2). Woolley (Woolley 1976; Woolley and Sutcliffe 1977; Woolley 1978a; Woolley 1978b; Woolley 1980) has argued with some justification that when experimental structure determinations reveal broken symmetry in molecules, this must be due to weak inter-molecular or other environmental perturbations. The wave function for the first excited state  $\Psi_-$  of ammonia is the corresponding antisymmetric linear combination:

$$\Psi_+ = (1/\sqrt{2}) (\Psi_L + \Psi_R) \quad (6)$$

$$\Psi_- = (1/\sqrt{2}) (\Psi_L - \Psi_R)$$

Indeed, excitation of a proton from  $\Psi_+$  to  $\Psi_-$  gives rise to a line in the protonic spectrum that corresponds to the observed microwave line in ammonia; this is conventionally explained as due to umbrella-flipping transformation between  $\Psi_L$  and  $\Psi_R$  (Fig. 2).

Thomas thus provides an attractive alternative formalism for obtaining — and visualizing — molecular structure without making the BOA, and treating electrons and protons on a similar footing. Such an approach has also recently been adopted by Nakai (2002) and coworkers (Nakai 2002; Nakai, Hoshino et al. 2005; Hoshino and Nakai 2006), with much greater computational rigor, in their NOMO (nuclear orbital plus molecular orbital) theory. This goes to show that there are many equally valid ways of conceptualizing molecular structure. While solution of the electronic Schrödinger equation within BOA takes the coordinates of the nuclei as input parameters, no such input is required for computing the eigenfunctions of the total molecular Hamiltonian (5), which only takes the numbers of each kind of particle as input. A solution of the total molecular Hamiltonian of  $C_6H_6$  will thus give all the structures in Fig.3, benzene as well as prismane, *etc.* If the structures are not well separated energetically, a single solution could have significant contributions from several of these structures. While at first sight this may seem like a nasty complication, it provides a natural treatment of resonance, tautomerism, Jahn-Teller molecules, molecular transformations and fluxional molecules.

Konarski (Konarski 1987; Konarski 1992; Konarski 1994) has used the term “soft body” to describe such situations. Early (Early 2005) used a similar argument to question why structural isomers are generally considered different chemical substances, while the different phases of water are recognized as different forms of one chemical substance. In

the approach described above, these forms and isomers would all be treated on an equal footing.

### **Jahn-Teller Molecules and the Geometric Phase:**

*“If the wave function of a given electronic state changes sign when transported adiabatically around a closed loop in nuclear configuration space, then the state must become degenerate with another at some point within the loop. This sign reversal condition is necessary and sufficient to establish the existence of an intersection.”*

- G. Herzberg & H. C. Longuet-Higgins (Herzberg and Longuet-Higgins 1963)

A remarkable consequence of the breakdown of BOA is the appearance of a geometric phase (Berry's phase) (Mead and Truhlar 1979; Berry 1984; Shapere and Wilczek 1989) in molecular systems. In 1963 Herzberg and Longuet-Higgins (Herzberg and Longuet-Higgins 1963; Longuet-Higgins 1975), predicted a change of sign of the electronic wave function around a closed loop of structure deformations encircling a Jahn-Teller (JT) intersection between two electronic states. This sign change is a manifestation of the geometric phase. Jahn-Teller molecules are chemical manifestations of broken symmetry; this symmetry breaking also lifts the degeneracy. A prototypical JT molecule is the system of three hydrogen atoms ( $H_3$ ) in its equilateral configuration. The equilateral geometry is unstable and either of the degenerate vibrational motions  $Q_1$  and  $Q_2$  (Fig.4) which break the molecular symmetry, result in a lowering of the energy, leading in this case to dissociation to diatomic  $H_2$  + H atom. In the unstable symmetrical structure, an infinitesimal perturbation is sufficient to set the electron density in motion. In QTAIM

this is described as a catastrophe point of the molecular graph and in density functional theory as a process of bond melting (Sukumar 1994; Sukumar 1995). The geometric phase can also be variously interpreted as a magnetic monopole gauge potential, a molecular manifestation of the Aharonov Bohm effect or a change in the effective action (4) accompanying variation in external parameters (nuclear coordinates). The predicted sign change (a molecular Aharonov Bohm effect) around the point of JT degeneracy was experimentally confirmed in 1986 by the observation of fractional quantization of molecular pseudorotation in  $\text{Na}_3$  (Delacretaz, Grant et al. 1986). The contributions to the phase due to permutation symmetry of identical nuclei predicted by Mead and Truhlar (Mead and Truhlar 1979) have also been experimentally confirmed. In periodic crystals, the geometric phase leads to macroscopic polarization (Ortiz and Martin 1994; Resta 1994; Gonze, Ghosez et al. 1997; Martin 1997) in the density functional theory of polar insulators. Resta (Resta 1998) has shown that the expectation value of the position operator — usually a “structural” parameter related to the molecular dipole moment — is now related to the geometric phase in periodic systems. So structure and phase are intimately related.

## **Hierarchy of protein structures:**

*“The hierarchy of approximations creates a path... to scientifically constructed qualitative images, notions and patterns that emerge from ‘exact’ equations.”*

— Valentin Ostrovsky (Ostrovsky 2005)

Conformational flexibility, as manifested in JT systems, is also a critical link between protein structure, stability and biological function. Enzymes must be flexible enough to mediate a reaction pathway, yet rigid enough to achieve molecular recognition. Proteins have a hierarchy of structures shown in Fig.5:

- Primary structure, specified by the sequence of amino acid residues;
- Secondary structure, specified by the dihedral angles  $\phi$  and  $\psi$  around the rotatable bonds
- Tertiary structure, referring to the steric relationships of amino acid residues that are far apart in the linear sequence;
- Quaternary structure, between individual subunits constituting the protein.

Furthermore, proteins are always associated with weakly bound water that critically influences their structure, dynamics and biological function, and providing the thermodynamic driving force for protein folding and self-assembly. It is thus no more legitimate to talk of proteins without water as it is to consider salt and water as independent entities in sea water (Early 2005).

Let us reflect on the notion of structure as introduced in the first section. If a protein were merely a heap of amino acid residues, we would expect protein structures to uniformly populate the  $\phi$ - $\psi$  plane. That this is far from the case can be seen from examination of the Ramachandran map (Fig.5), which reveals correlations between the dihedral angles  $\phi$  and  $\psi$ . Different regions of the Ramachandran map correspond to qualitatively distinct structural motifs, such as the  $\alpha$ -helix, the  $\beta$ -sheet and the left-handed helix. In fact, higher-order correlations have also been discovered, extending across dihedral angles

spanning multiple residues (Sims and Kim 2005). Similarly Mezei (Mezei 2003), in accord with the statistical definition of structure, introduced a fingerprint for the characterization of protein folds, in terms of the angular correlations between the C=O direction, the backbone axis and the line connecting the alpha carbons of residues.

The complexity of protein conformational landscapes makes straightforward energy minimization of a protein to its global energy minimum (the “native structure”) a hopelessly demanding computational task. Traditional approaches to this problem have exploited the analogy with spin glasses and employed computational and statistical techniques such as simulated annealing and Monte Carlo sampling. But proteins in nature seem to “know” how to fold to their native structures: these being determined from the sequence of amino acid residues or the primary structure. Solving this puzzle — understanding the mechanism involved and devising computational algorithms to simulate the process — is the protein-folding problem. Recent progress in solving this puzzle reveals that many proteins seem to be well described by a funnel-like free-energy landscape, with often no apparent microscopic energetic or entropic barriers (Schonbrun and Dill 2003). The relatively fast timescale of the folding process seems to exploit massive parallelism of multiple paths or folding trajectories on this surface. This then is a very different view of molecular structure from the classical and connects with the systems or process view of chemistry (Reiher 2003; Stein 2006), which gives ontological priority to becoming over being and emphasizes molecular transformations, represented by free-energy landscapes, as the primary currency of chemistry.

## **Molecular structure and Molecular similarity:**

*“Hence there is probably no wonder that, even for very small molecules, it has been found that different experimental methods may lead to different molecular structures, indicating that no such unique structure exists.”*

— P.-O. Löwdin (Löwdin 1991)

While philosophers deal primarily with the meaning of concepts, physical scientists are always concerned with quantification thereof. Quantification of molecular structure is the domain of molecular similarity analysis (Carbo-Dorca, Robert et al. 2000; Bender, Jenkins et al. 2006; Agrafiotis, Bandyopadhyay et al. 2007). Various similarity measures derived from molecular structure have been employed to classify and cluster molecules in pharmaceutical and public-domain compound collections, some of which work directly on molecular graphs (Batista and Bajorath 2007) and others on shape (Breneman, Sundling et al. 2003; Zauhar, Moyna et al. 2003). Since molecular structure is what distinguishes a molecule from a collection of its constituent atoms, the properties of a molecule must be different from those of a heap of its constituent atoms and must depend on the molecular structure. Quantification of the relationship between molecular structure and properties goes by the name of Quantitative Structure Property Relationship (QSPR) modeling (Hansch, Muir et al. 1963; Agrafiotis, Bandyopadhyay et al. 2007), a very active field of research for nearly half a century. But as there is often no direct mathematical relationship between molecular structure and molecular property, an intermediary is required in the form of molecular descriptors — numerical representations of molecular structure, in a form tractable for statistical modeling.

Relationships between descriptors and properties are often those of correlation, not causation. Zeidler (Zeidler 2000), following Giere (Giere 1994), distinguishes between two kinds of representation: structural (which implies a structural similarity between representing and represented objects) and informational (where the representing object is merely a vehicle of some information pertaining to the represented object). Descriptors are thus clearly in the latter category. Further, different types of descriptors are effective in modeling the relationship to different properties. For instance, properties like boiling point, non-specific toxicity and passive transport (permeability) across biological membranes are often adequately modeled by topological descriptors, derived from the molecular graph alone, whereas descriptors derived from and dependent upon the 3-D structure and conformation are generally necessary to model binding to specific protein receptor sites.

However, since structural parameters depend upon the experimental procedures employed in their determination (Löwdin 1991), Zeidler (Zeidler 2000), goes on to argue that dynamic models of molecular structure (*i.e.* of the kind implied by Konarski's (Konarski 1987; Konarski 1992; Konarski 1994) "soft bodies" or indeed, anything beyond the classical ball-and-stick kind) cannot be regarded as structural models, but merely as informational ones. But if we accept this argument, there is no reason to regard molecular descriptors, which are quantitative, numerical, representations of a molecule based on its structure, in any different light philosophically, than other informational models of molecular "structure". This argument may be examined in light of the realization (Brown 2003a; Brown 2003b) that descriptors and models of molecular structure, in terms of which chemists conceptualize the chemical universe, are metaphors

employed to relate the domain of chemical phenomena in terms of the everyday known and familiar (Lakoff 1993).

## **Conclusion:**

We have examined molecular structure from a generalized statistical viewpoint and argued that conventional notions of molecular structure are a limiting case in a hierarchy of structural conceptions. The conventional view of molecular structure arises from the BOA, although it is possible to define structure without specific reference to BOA. We discussed molecular structure in the QTAIM and showed how the approximate transferability of atomic and functional group properties among similar molecules in this theory enables the systematic study of molecular properties that is at the heart of chemistry. The consequences of going beyond the BOA were analyzed and illustrated for the example of JT molecules. Protein structure was examined in light of insights from these conceptions of molecular structure. We were thus led to the process view of protein chemistry (Stein 2006) and the conception of the free energy landscape as the primary structural entity. Finally we examined the quantification of the relationship between molecular structure and properties, through the intermediary of mathematical descriptors, which were regarded as informational models of molecular structure. Since the models in terms of which science is formulated and the metaphors relating the domain of study to the domain of parlance directly affect the explanatory and interpretive power of chemistry (Brown 2003a; Brown 2003b), we make the case for taking a broad view of molecular structure, emphasizing the fresh insights to be gained from different complementary models.

## Figure Captions

Figure 1: (a) Electron density contours and (b) gradient paths ( $\nabla\rho$ ) of tetrahedrane, showing the bond paths and interatomic surfaces; (c) electron density contours and (d) gradient paths ( $\nabla\rho$ ) of benzene in the molecular plane.

Figure 2: The structure and protonic spectrum of ammonia. The lower part of the figure shows the radial distribution functions for the protons as functions of the N-H distance.

Figure 3: What is the structure of  $C_6H_6$ ? All these structures are described by the *same* molecular Hamiltonian:  $H(6 \text{ carbons}, 6 \text{ hydrogens}, 42 \text{ electron soup})$

Figure 4: (a) Jahn-Teller vibrational modes  $Q_1$  and  $Q_2$  and (b) electronic energy levels along one of the modes of a prototypical Jahn-Teller molecule  $H_3$ . The equilateral geometry is unstable and either of the modes breaks the molecular symmetry and lower the ground state energy. Adapted from (Sukumar 1984).

Figure 5: Hierarchy of protein structures showing (a)  $\alpha$ -helix and  $\beta$ -sheet structures (b) Ramachandran map for turkey egg white lysozyme and (c) definitions of the dihedral angles  $\varphi$ ,  $\psi$  and  $\omega$ .

## References

- Ackeroyd, M.: Explanation of the Atomic Theory 1865-1895. International Society for the Philosophy of Chemistry Summer Symposium, Knoxville, TN (2005).
- Agrafiotis, D. K., Bandyopadhyay, D., et al.: Recent advances in cheminformatics. *J. Chem. Inf. Model.* **47**(4): 1279-1293 (2007).
- Bader, R. F. W.: Quantum topology of molecular charge distributions. III. The Mechanics of an Atom in a Molecule. *Journal of Chemical Physics* **73**(6): 2871-2883 (1980).
- Bader, R. F. W.: *Atoms in Molecules: A Quantum Theory*. Clarendon, Oxford (1990).
- Bader, R. F. W.: A quantum theory of molecular structure and its applications. *Chem. Rev.* **91**(5): 893-928 (1991).
- Bader, R. F. W.: *Atoms In Molecules: A Quantum Theory*. Clarendon, Oxford (1995).
- Bader, R. F. W.: 1997 Polanyi Award Lecture. Why are there atoms in chemistry? *Can. J. Chem.* **76**(7): 973-988 (1998).
- Bader, R. F. W.: A Bond Path: A Universal Indicator of Bonded Interactions. *J. Phys. Chem. A* **102**(37): 7314-7323 (1998).
- Bader, R. F. W., Anderson, S. G., et al.: Quantum topology of molecular charge distributions. 1. *J. Am. Chem. Soc.* **101**(6): 1389-95 (1979).
- Bader, R. F. W., Nguyen-Dang, T. T., et al.: Quantum topology of molecular charge distributions. II. Molecular structure and its change. *J. Chem. Phys.* **70**: 4316-4329 (1979).
- Bader, R. F. W., Nguyen-Dang, T. T., et al.: A topological theory of molecular structure. *Rep. Prog. Phys.* **44**: 893-948 (1981).
- Bader, R. F. W., Popelier, P. L. A., et al.: Theoretical Definition of a Functional Group and the Molecular Orbital Paradigm. *Angew Chem. Int. Ed. Engl* **33**: 620-631 (1994).
- Bader, R. F. W., Srebrenik, S., et al.: Subspace quantum dynamics and the quantum action principle. *J. Chem. Phys.* **68**(8): 3580-91 (1978).
- Batista, J., Bajorath, J.: Chemical database mining through entropy-based molecular similarity assessment of randomly generated structural fragment populations. *J. Chem. Inf. Model.* **47**: 59-68 (2007).
- Bender, A., Jenkins, J. L., et al.: Molecular Similarity: Advances in Methods, Applications, and Validations in Virtual Screening and QSAR. *Ann. Rep. Comp. Chem.* **2**: 141 - 168 (2006).

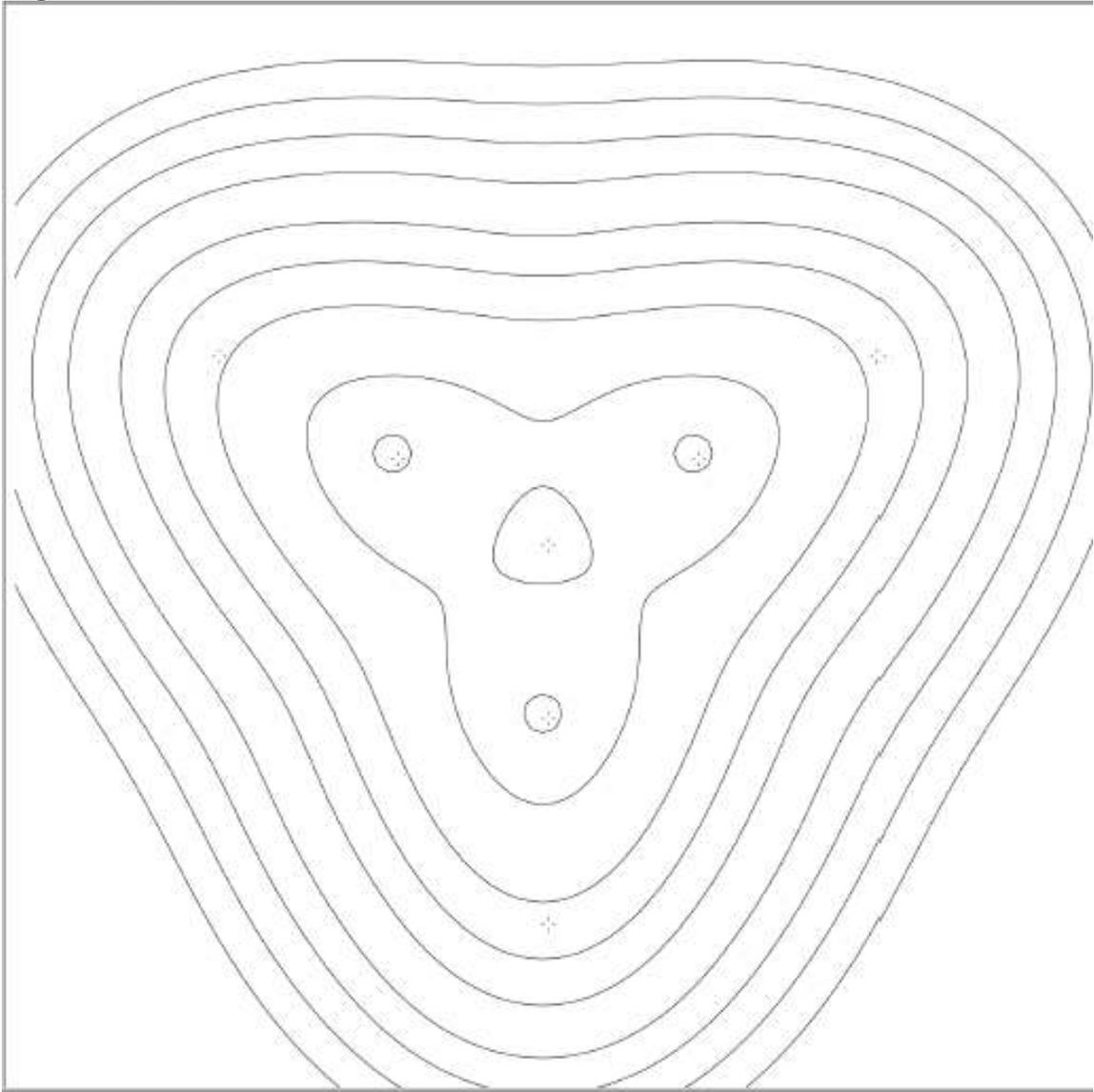
- Berry, M. V.: Quantal Phase Factors Accompanying Adiabatic Changes. Proc. Roy. Soc. A **392**: 45-57 (1984).
- Born, M.: Nachr. Akad. Goettingen Math.-Physik.Kl. **6**: 1 (1951).
- Born, M., Huang, K.: Dynamical Theory of Crystal Lattices. Clarendon, Oxford (1954).
- Born, M., Oppenheimer, J. R.: Ann. Physik **84**: 457 (1927).
- Breneman, C. M., Sundling, C. M., et al.: New developments in PEST shape/property hybrid descriptors. J. Comput. aided Mol. Design **17**: 231-240 (2003).
- Brown, T. L.: Making Truth: Metaphor in Science. Univ. Illinois Press, Urbana (2003a).
- Brown, T. L.: The Metaphorical Foundations of Chemical Explanation. Ann. NY Acad. Sci. **988**: 209-216 (2003b).
- Carbo-Dorca, R., Robert, D., et al.: Molecular Quantum Similarity in QSAR and Drug Design. Springer-Verlag, Berlin (2000).
- Coulson, C. A. Tilden Lecture, Burlington House, The Chemical Society (1951).
- Coulson, C. A.: The contributions of wave mechanics to chemistry. J. Chem. Soc.: 2069-2084 (1955).
- Delacretaz, G., Grant, E. R., et al.: Fractional Quantization of Molecular Pseudorotation in Na<sub>3</sub>. Phys. Rev. Lett. **56**: 2598-2601 (1986).
- Early, J. E.: Why there is no salt in the sea. Found. Chem. **7**: 85-102 (2005).
- Eberhart, M.: Quantum Mechanics and Molecular Design in the Twenty-First Century. Found. Chem. **4**: 201-211 (2002).
- Essén, H.: The physics of the Born-Oppenheimer approximation. Int. J. Quantum Chem. **12**: 721-735 (1977).
- Essén, H. Energy Storage and Redistribution in Molecules. J. Hinze: pp. 327 Plenum, New York, (1983).
- Giere, R. N.: No representation without representation. Biology and Philosophy **9**: 113-120 (1994).
- Gonze, X., Ghosez, P., et al.: Density functional theory of polar insulators. Phys. Rev. Lett. **78**(2): 294-297 (1997).
- Hansch, C., Muir, R. M., et al.: The correlation of biological activity of plant growth regulators and chloromycetin derivatives with Hammett constants and partition coefficients. J. Amer. Chem. Soc. **85**: 2817-2824 (1963).

- Herzberg, G., Longuet-Higgins, H. C.: Intersection of potential energy surfaces in polyatomic molecules. *Discuss. Faraday Soc.* **35**: 77-82 (1963).
- Hohenberg, P., Kohn, W.: Inhomogeneous Electron Gas. *Phys. Rev.* **136**: B864-B871 (1964).
- Hoshino, M., Nakai, H.: Elimination of translational and rotational motions in nuclear orbital plus molecular orbital theory: Application of Møller-Plesset perturbation theory. *J. Chem. Phys.* **124**: 194110 (2006).
- Konarski, J.: Diatomic molecule as a soft body. *J. Molec. Spectry.* **124**(1): 218-228 (1987).
- Konarski, J.: Rovibrational states of a linear molecule. *J. Molec. Struct.* **270**: 491-498 (1992).
- Konarski, J.: A new model of a molecule based on the soft body. *Int. J. Quantum Chem.* **51**: 439-445 (1994).
- Lakoff, G.: The contemporary theory of metaphor. *Metaphor and Thought*. A. Ortony: Cambridge Univ. Press, Cambridge, (1993).
- Lathouwers, L., VanLeuven, P.: Molecular spectra and the generator coordinate method. *Int. J. Quantum Chem.* **12S**: 371-375 (1978).
- Lathouwers, L., VanLeuven, P., et al.: Quantum theory and molecular spectra. *Chem. Phys. Lett.* **52**(3): 439-441 (1977).
- Longuet-Higgins, H. C.: The intersection of potential energy surfaces in polyatomic molecules. *Proc. Roy. Soc. A* **344**: 147 (1975).
- Löwdin, P.-O.: On nuclear motion and the definition of molecular structure. *J. Molec. Struct. (THEOCHEM)* **230**: 13-15 (1991).
- Martin, R. M.: Functional theory of extended Coulomb systems. *Phys. Rev. B* **56**(3): 1124-1140 (1997).
- Matta, C., Boyd, R. J., eds): *The Quantum Theory of Atoms in Molecules: From Solid State to DNA and Drug Design*. Wiley-VCH (2007).
- Mead, C. A., Truhlar, D. G.: On the determination of Born-Oppenheimer nuclear motion wave functions including complications due to conical intersections and identical nuclei. *J. Chem. Phys.* **70**: 2284-2296 (1979).
- Mezei, M.: A novel Fingerprint for the characterization of protein folds. *Protein Eng.* **16**: 713-715 (2003).

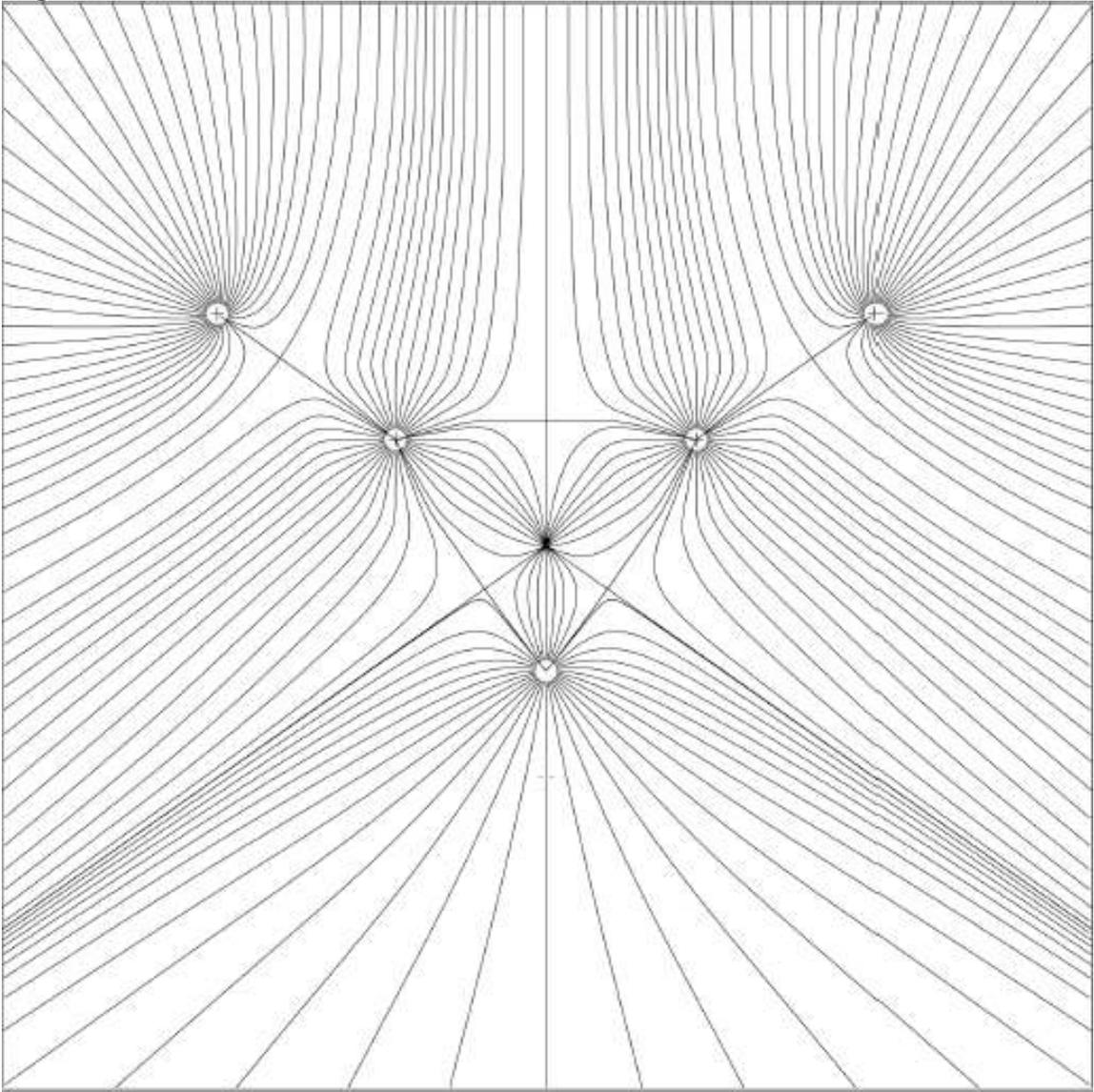
- Nakai, H.: Simultaneous Determination of Nuclear and Electronic Wave Functions without Born-Oppenheimer Approximation: Ab Initio NO+MO/HF Theory. *Int. J. Quantum Chem.* **86**: 511-517 (2002).
- Nakai, H., Hoshino, M., et al.: Elimination of translational and rotational motions in nuclear orbital plus molecular orbital theory. *J. Chem. Phys.* **122**: 164101 (2005).
- Ortiz, G., Martin, R. M.: Macroscopic polarization as a geometric quantum phase: Many-body formalism. *Phys. Rev. B* **49**(20): 14202-14210 (1994).
- Ostrovsky, V.: Towards a Philosophy of Approximations in the 'Exact' Sciences. *HYLE* **11**(2): 101-126 (2005).
- Reiher, M.: The Systems-Theoretical View of Chemical Concepts. *Found. Chem.* **5**: 147-163 (2003).
- Resta, R.: Macroscopic polarization in crystalline dielectrics: the geometric phase approach. *Rev. Mod. Phys.* **66**(3): 899-915 (1994).
- Resta, R.: Quantum-Mechanical Position Operator in Extended Systems. *Phys. Rev. Lett.* **80**: 1800 - 1803 (1998).
- Schonbrun, J., Dill, K. A.: Fast protein folding kinetics. *Proc. Natl. Acad. Sci. U. S. A.* **100**(22): 12678-12682 (2003).
- Schwinger, J.: The theory of quantized fields. I. *Phys. Rev.* **82**: 914 (1951).
- Shahbazian, S., Zahedi, M.: The Role of Observables and Non-Observables in Chemistry: A Critique of Chemical Language. *Found. Chem.* **8**: 37-52 (2006).
- Shapere, A., Wilczek, F.: *Geometric Phases in Physics*. World Scientific, Singapore (1989).
- Sims, G. E., Kim, S.-H.: A method for evaluating the structural quality of protein models by using higher-order  $\phi$ - $\psi$  pairs scoring. *Proc. Natl. Acad. Sci. U. S. A.* **102**(12): 618-621 (2005).
- Smith, C. S.: *A Search for Structure*. MIT Press, Cambridge (1981).
- Stein, R. L.: A Process Theory of Enzyme Catalytic Power - The Interplay of Science and Metaphysics. *Found. Chem.* **8**: 3-29 (2006).
- Sukumar, N.: *Born Couplings in H<sub>2</sub><sup>+</sup>, H<sub>2</sub> and H<sub>3</sub>. Chemistry*. Stony Brook, State University of Stony Brook. Ph. D. (1984).
- Sukumar, N.: Density Functional Theory for Jahn-Teller Systems. *Int. J. Quantum Chem.* **52**: 809-816 (1994).

- Sukumar, N.: Density Functional Theory of Born Couplings. *Int. J. Quantum Chem.* **56**: 423-432 (1995).
- Sukumar, N., Breneman, C. M.: QTAIM in Drug Discovery. *Quantum Theory of Atoms in Molecules*. C. Matta and D. Boyd: Wiley-VCH, (2006).
- Sutcliffe, B. T.: The chemical bond and molecular structure. *J. Molec. Struct. (THEOCHEM)* **259**: 29-58 (1992).
- Thomas, I. L.: Protonic Structure of Molecules. I. Ammonia Molecules. *Phys. Rev.* **185**: 90 (1969).
- Thomas, I. L.: Selection Rules and the Protonic Spectrum of Molecules. *Phys. Rev. A* **2**: 72 (1970).
- Thomas, I. L.: Stark and Zeeman Effects on the Protonic Structure of Molecules. *Phys. Rev. A* **2**: 1675 (1970).
- Thomas, I. L.: Vibrational and Rotational Energy Levels as Protonic Structure in Molecules. *Phys. Rev. A* **3**: 565 (1971).
- Thomas, I. L.: Photoprotonic Effect in Hydrides. *Phys. Rev. A* **5**: 1972 (1972).
- Thomas, I. L., Joy, H. W.: Protonic Structure of Molecules. II. Methodology, Center-of-Mass Transformation and the Structure of Methane, Ammonia and Water. *Phys. Rev. A* **2**: 1200 (1970).
- Trindle, C.: The Quantum Mechanical View of Molecular Structure and the Shapes of Molecules. *Israel J. Chem.* **19**: 47-53 (1980).
- Woolley, R. G.: Quantum theory and molecular structure. *Adv. Phys.* **25**: 27-52 (1976).
- Woolley, R. G.: Further remarks on molecular structure in quantum theory. *Chem. Phys. Lett.* **55**: 443-446 (1978a).
- Woolley, R. G.: Must a molecule have a shape? *J. Amer. Chem. Soc.* **100**: 1073-1078 (1978b).
- Woolley, R. G.: Quantum Mechanical Aspects of the Molecular Structure Hypothesis. *Israel J. Chem.* **19**: 30-46 (1980).
- Woolley, R. G., Sutcliffe, B. T.: Molecular structure and the Born-Oppenheimer approximation. *Chem. Phys. Lett.* **45**(2): 393-398 (1977).
- Zauhar, R., Moyna, G., et al.: Shape Signatures: a new approach to computer-aided ligand- and receptor-based drug design. *J. Med. Chem.* **46**: 5674-5690 (2003).
- Zeidler, P.: The epistemological status of theoretical models of molecular structure. *HYLE* **6**(1): 17-34 (2000).

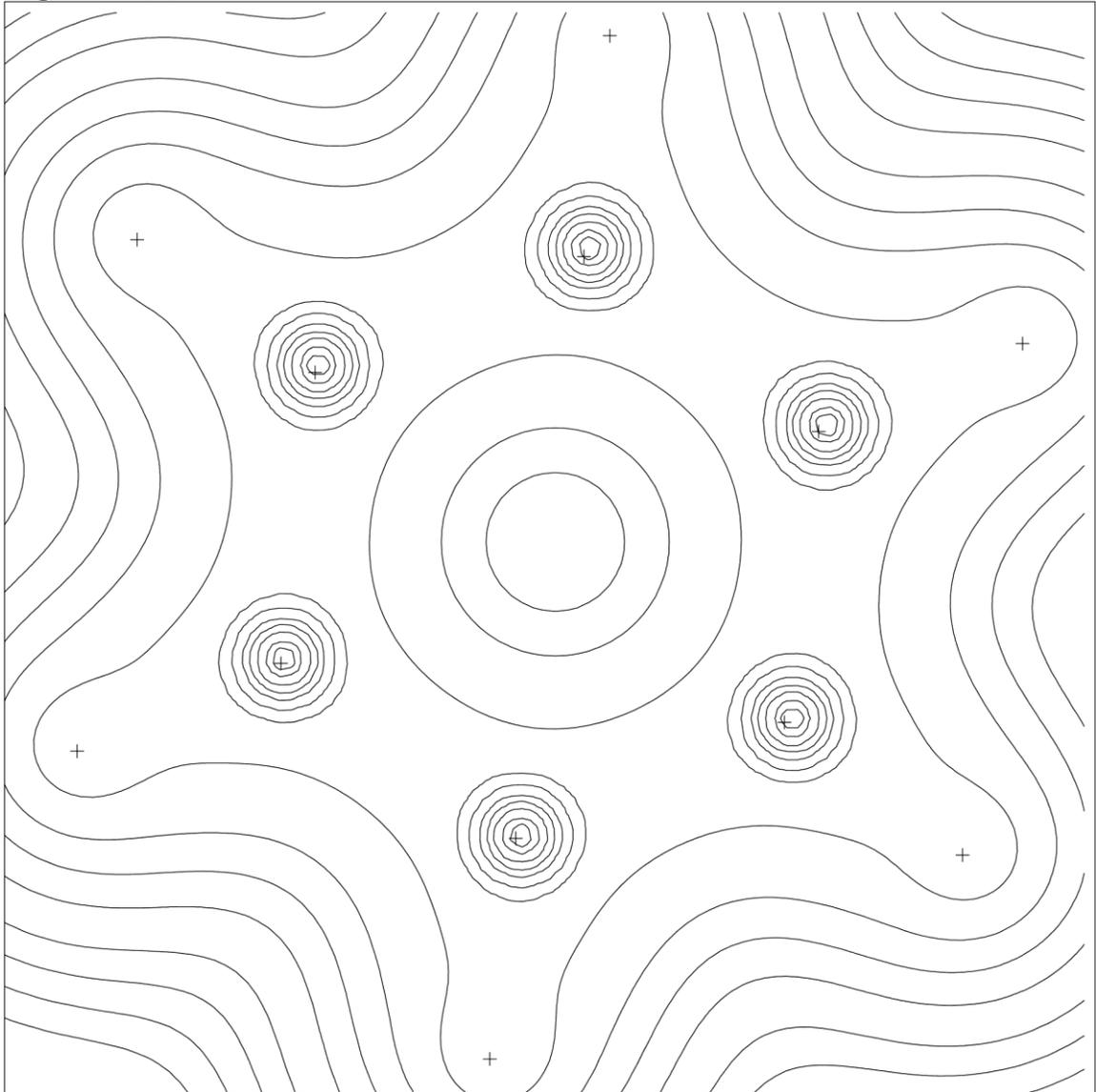
**Figure 1 (a)**



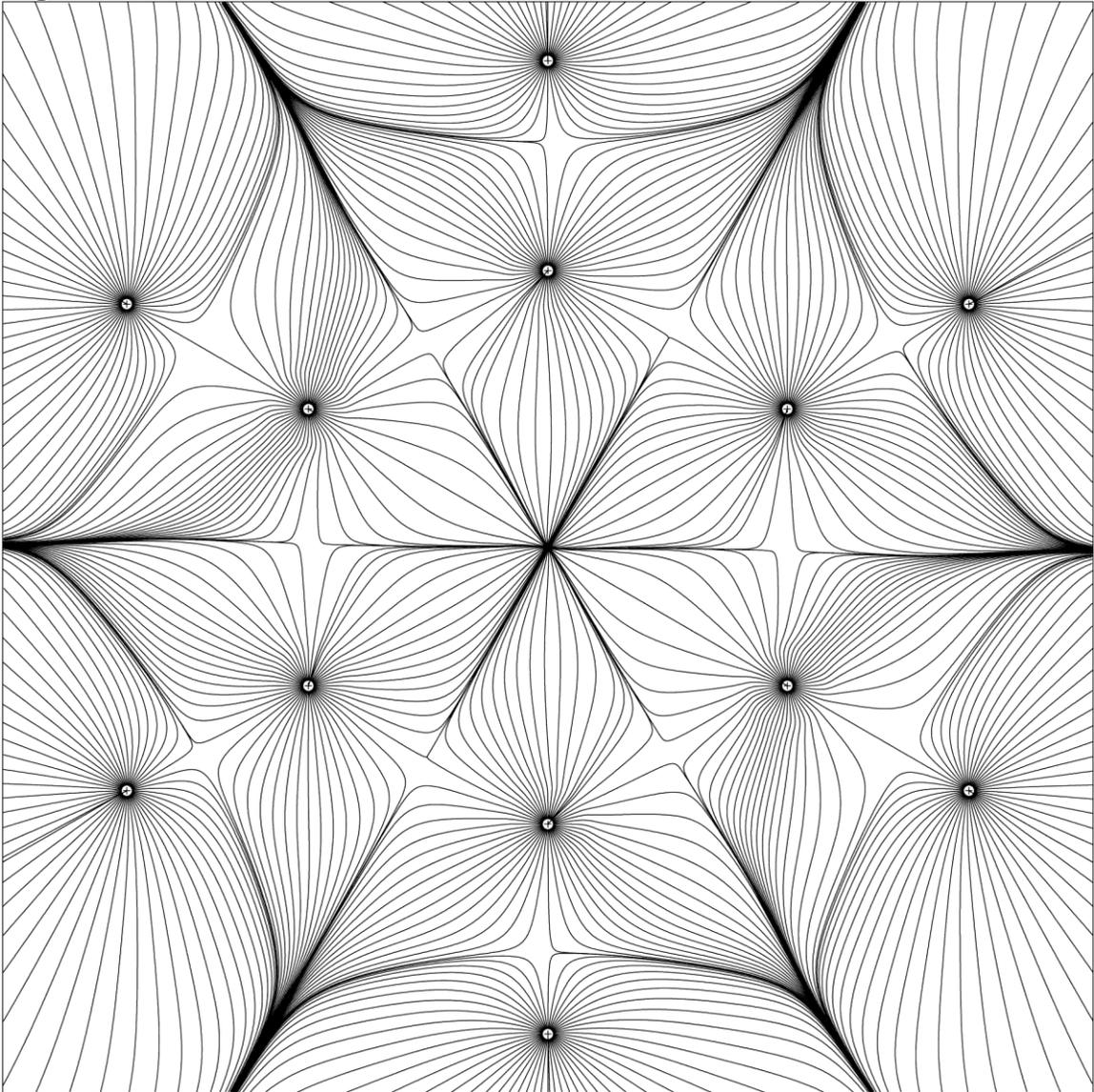
**Figure 1 (b)**



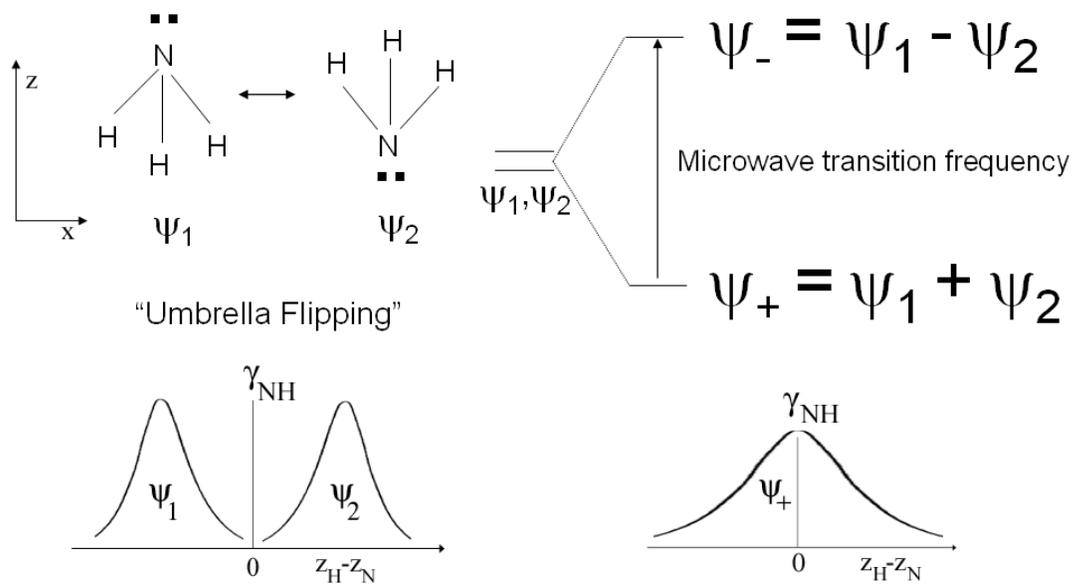
**Figure 1 (c)**



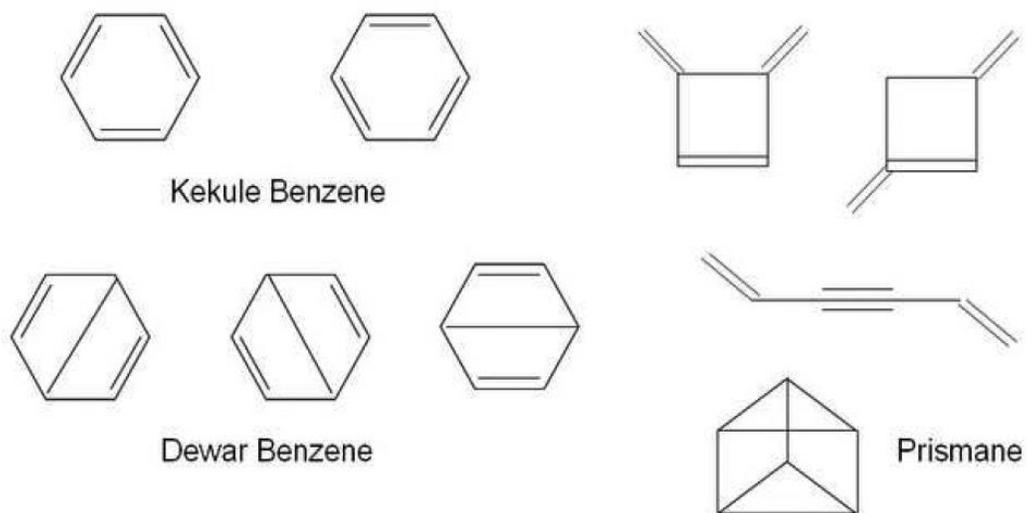
**Figure 1 (d)**



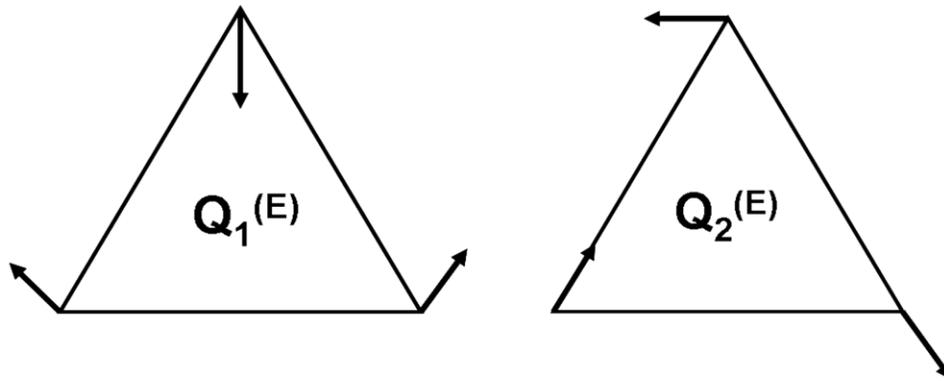
**Figure 2**



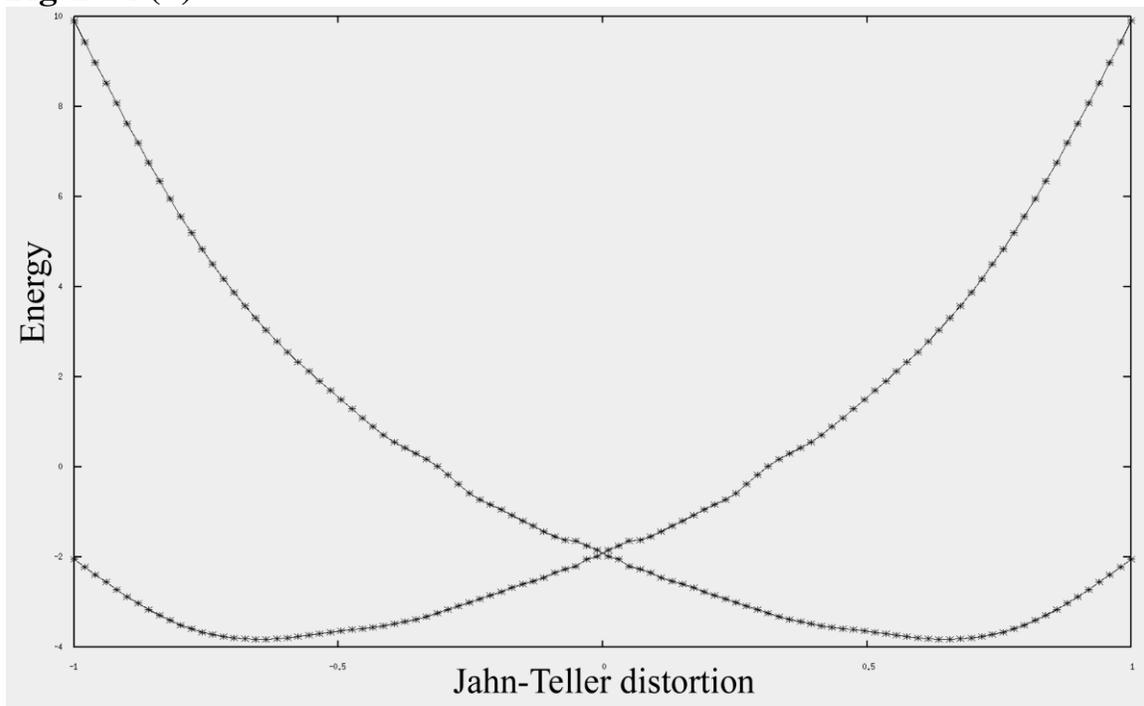
**Figure 3**



**Figure 4 (a)**



**Figure 4 (b)**



**Figure 5 (a)**

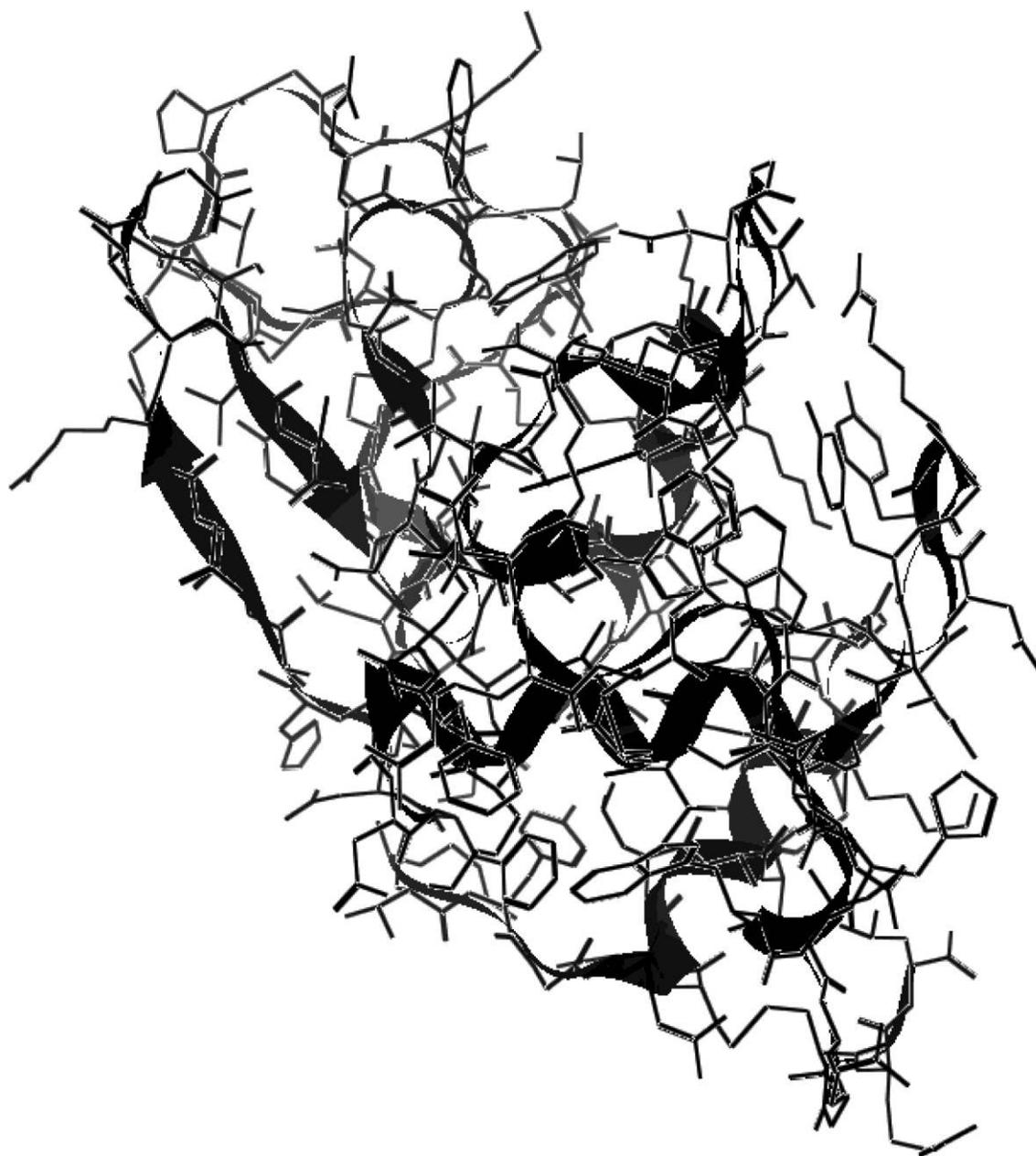


Figure 5 (b)

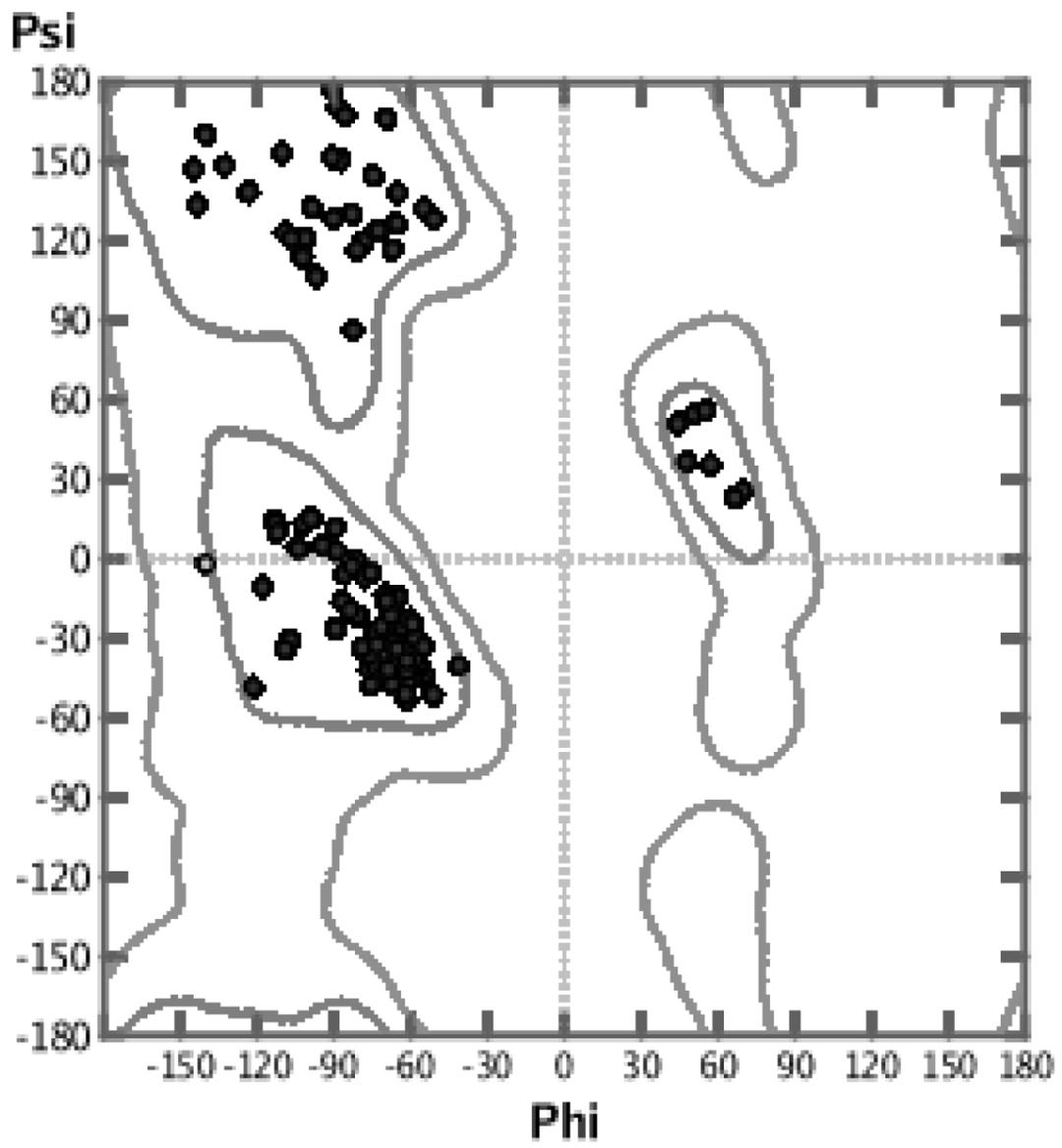


Figure 5 (c)

