ELSEVIER

Coordination Chemistry Reviews
212 (2001) 3–10

COORDINATION CHEMISTRY REVIEWS

www.elsevier.com/locate/ccr

Molecular mechanics: theoretical basis, rules, scope and limits

Jan C.A. Boeyens* [a,1], Peter Comba* [b,2]

[a] *Department of Chemistry, University of the Witwatersrand, 1 Jan Smuts Avenue, Johannesburg 2001, South Africa*
[b] *Universität Heidelberg, Anorganisch-Chemisches Institut, Im Neuenheimer Feld 270, 69120 Heidelberg, Germany*

Received 3 September 1999; received in revised form 26 January 2000; accepted 10 March 2000

Contents

[1] *Corresponding author. Tel.: +49-6221-548453; fax: +49-6221-546617, e-mail: j.boeyens@postino.up.ac.za

[2] *Corresponding author. Tel.: +49-6221-548453; fax: +49-6221-546617, e-mail: comba@akcomba.oci.uni-heidelberg.de

PII: S0010-8545(00)00353-2

Abstract

Molecular mechanics is a simple model, and it is based on a classical parameterization of non-classical effects for the computation of molecular structure. The basic concept of force field calculations is discussed and, based on the theoretical frame, the emerging scope and limits of the approach, and rules for the application, interpretation of the results and their communication are presented.

Keywords: Molecular mechanics; Force field; Strain energy; Structure; Molecular modeling

1. Introduction

A molecular structure is the three-dimensional arrangement of the atoms of a molecule in space. The basic interest in the description of classical and static molecular structures is the desire for understanding by visualization and the belief that molecular properties (thermodynamics, i.e., stabilities; kinetics, i.e., reactivities; electronics, i.e., spectroscopy) are directly related to molecular structure. Among the various methods that are used to compute structures molecular mechanics (MM) has been and still is appreciated as an approach that is able to optimize structures with high accuracy and comparably little computational expense. The exponential growth of low cost computing power, the optimization of computational tools and the development of new and increasingly efficient quantum mechanical methods do not necessarily lead to a loss of interest in MM-based approaches, especially, when a fast and accurate computation of structures of appreciably large molecules, an extensive conformational search, dynamic processes and/or the interaction between isolated molecules and the environment are the aim of a study.

MM has been reviewed extensively in recent years [1–24]. However, so far there is no short, fundamental, comprehensive and clear summary on what MM exactly is, how and where it can be used, how results should be reported and interpreted, and where the limits, pitfalls and dangers are. Therefore, in this note, we will not comment on specific applications, on functional forms, on parameterization schemes or computational procedures but only on the generally accepted concept and the ensuing possibilities and problems.

2. The concept

2.1. The chemical bond

The chemical bond between a pair of atoms is a function of the electron density distribution between the atoms. In MM this function is quantified in terms of the two atom types which usually also define the bond order N (e.g. C(sp), C(sp), $N = 3$). Given atom types (and bond orders) imply a specific equilibrium (or ideal)

bond distance (r_0) and a specific force that needs to be applied to distort (stretch or compress) this bond (force constant k_b). In the simplest case, the potential which describes such a distortion is harmonic ($E_b = 1/2k_b(r - r_0)^2$, Hookes law). Steric effects (enforced distortions of the bond by other atoms in the molecule or by the environment) do not affect the bond order. Thus, the parameters k_b and r_0 are only dependent on the atom types, and they are fully transferable into any steric environment.

The atom types are assigned on the basis of the connectivity in a molecule, that is, different bonding parameter sets (k_b; r_0) and, therefore, different atom types have to be defined for example for carbon–carbon single bonds that involve a methyl group on one side and a carbonyl group, an aromatic carbon atom, a carboxylate or an aliphatic carbon atom on the other; also, different atom types and bonding parameters (k_b; r_0) are needed for nickel–amine nitrogen bonds, when nickel is in the +I, +II or +III oxidation state, when it is four-, five- or six-coordinate, when it is square planar or tetrahedral etc. That is, information on electronic factors is included in the assignment of the atom types. Molecular mechanics is a classical parameterization of non-classical effects.

The observation that characteristic parameters (e.g. r_0) may be sensitive to intramolecular electronic factors indicates that integral bond orders are not necessarily the rule. Even in the absence of steric factors and substituent effects the possible bonding modes for a given pair of atoms may define a continuous space where each point represents a specific parameter set r_0, k_b and N, where N is the bond order. This is, for example, the case in metal–metal bonds, where r_0, k_b and N have been found to be related by the expression $b \cdot N = k_b = c r_0^{-a}$, where a, b, and c are characteristic parameters for a specific pair of metal centers [25]. Parameter values may be obtained by the computation of known molecular structures which have equivalent bonds in different steric environments [26].

2.2. Other terms

Arguments similar to those discussed above apply to all other terms that may describe electronic interactions in a molecular assembly. These include valence angles, torsional angles, the planarity of substructures (e.g. carboxylates or aromatic systems) and more specific distortion modes such as the tetrahedral twist in square planar, the trigonal twist in octahedral and the Berry twist in five-coordinate transition metal coordination compounds, Jahn–Teller distortion modes and *trans*-influences.

In addition to these electronic terms there are attractive and repulsive through-space interactions, and these include van-der-Waals interactions and electrostatic forces. Hydrogen-bonding is a separate term and, in some MM models, hydrogen-bonding potentials are used which involve some directionality. In a classical mechanical model it is inappropriate to discuss whether or not there is a theoretical basis and/or a necessity for this. Through-space interactions in general are of importance for intra- and inter-molecular interactions.

2.3. The result of a structure optimization

The sum of all energies that result from all the potentials that describe the interactions of each atom with each other in a molecular assembly (two,- three- and more-atom interactions; through-bond and through-space interactions) is the total strain energy (E_{total}). The force field (i.e. the ensemble of potential energy functions and their parameterization) allows the computation of the total strain energy at each point of the energy surface of a molecule or molecular assembly (i.e. for each set of coordinates), and strain energy minimization algorithms allow one to find energy minima by adjusting the coordinates of each atom. Thus, the results of an MM computation are an optimized structure and the corresponding minimized strain energy. Depending on the minimization algorithm used, vibrational frequencies are also available.

3. A general fact of importance

The central problem in chemistry is to interpret, understand and predict observable properties of chemical compounds: stabilities, reactivities, colors, solubilities etc. This involves understanding of observable bulk properties on a not observable molecular or submolecular level. This is only possible on the basis of molecular models, and these, very generally, are based on imagination (and, in many models, thorough and experimentally validated theories) rather than truth [27]. Molecular models are generally based on molecular structure which by itself is a non-trivial concept, and this has been discussed extensively [28,29]. MM is a simple model that is based on a classical parameterization of non-classical effects for the computation of molecular structure. It follows that the type of functions used, the degree of parameterization and the parameter values are only important in terms of the quality of agreement between computed and experimentally observed (structural) data. Indeed, different force fields which may lead to predictions of similar quality may use strikingly different sets of functions and parameterization schemes [7,13]. Thus, the functions and parameter values by themselves are not necessarily scientifically meaningful.

4. The force field

4.1. Potential energy functions

The classical MM concept assumes that individual potentials (e.g. bonding and angle bending terms) are independent. The neglect of coupling between the isolated terms does not necessarily lead to serious problems when the corresponding force constants are significantly different. Some force fields use cross-terms to improve the accuracy when the coupling is appreciable (non-diagonal force fields).

In its original form MM is based on harmonic potentials. For bond stretching, for example, this is only a valid assumption close to the potential energy minimum of each individual bond, and bond dissociation may not be modeled with harmonic potentials (nor with most anharmonic functions, except for the Morse potential).

A variety of functions have been used to model valence angle bending, in particular around transition metal coordination centers, where ligand field and other electronic factors are of importance. These have lead to highly sophisticated force fields that allow the accurate prediction of structures that involve open-shell compounds [30–34].

4.2. Force field parameters

Force field parameters for vibrational spectroscopy are not the same as those for MM calculations. While spectroscopic molecule-specific parameters describe small vibrations (distortions) around a molecule-specific equilibrium geometry that, for highly strained molecules, is far off the ideal structure, molecule-independent MM parameters may describe a relatively large deviation from a molecule-independent ideal structure. For this and other reasons are spectroscopic, molecule-specific force constants, obtained from normal coordinate analyses, at best starting values for fitting MM force field parameters to extensive sets of experimental data [13]. These may include crystal structures, thermodynamic data (heats of formation) and/or vibrational spectra. The resulting force field may then be used to compute corresponding data of similar molecules. Thus, MM is an interpolative method and extrapolation may lead to serious errors. Force fields have often been fitted to structural parameters. This is in particular true for force fields for transition metal coordination compounds, where accurate and thoroughly interpreted thermodynamic and vibrational spectroscopic data are not generally available. While structures relate to potential energy minima, relative stabilities are related to the steepness of the potentials and vibrations to the curvature.

Force field parameters have been obtained by calculation from first principles (rule-base generic force fields) [33,35,36], or, based on experimental data sets, parameters have been fitted with various optimization routines [37–39] and with genetic algorithms [40], but generally they still are obtained by manual fitting to experimentally observed structures. Potential energy functions and parameter sets within a force field are interdependent, that is, modifications to an existing force field require the whole force field to be checked carefully and parameters are not transferable from a validated force field to another.

4.3. Charges

Electrostatic interactions are a particular problem because there are only few methods for the accurate and quick computation of charges, in particular for inorganic compounds [13]. For relatively small molecules this usually is not a serious problem since Coulomb forces may be included in other potentials (bond stretching, valence angle bending, van-der-Waals interactions). However, when

intermolecular interactions are considered and for macromolecules Coulomb terms are relevant since their decrease with distance is considerably slower than for other functions.

5. Possibilities and problems

5.1. Structures

Usually, the optimized structure is a local minimum close to the starting structure. The search for the global energy minimum structure and the analysis of the conformational space must involve scanning of the energy surface. Methods available include grid search routines, molecular dynamics and stochastic (Monte Carlo) methods [13,41].

To prevent translation and rotation of a molecule during refinement it is common practice to constrain one atom of the molecule (x, y and z-coordinate), to constrain y and z of a second and z of a third atom. This may cause a problem if the first two of these atoms are linked, because their minimum separation $d_{min} = [\Delta y(1,2)^2 + \Delta z(1,2)^2]^{1/2}$ restricts the range of their allowed equilibrium distance (a value of d less than d_{min} is inaccessible) [42].

5.2. Strain energies

There is some thermodynamic information in strain energies. However, strictly this is only true if entropic terms are included, and this is not often the case. There are examples where the neglect of entropy does not lead to serious problems but this has to be evaluated carefully in each example. Also, strain energies are relative energies and the absolute energy values obtained in a structure optimization are meaningless. It emerges that only energy differences between isomers of a compound may be interpreted. The most serious restriction, however, is that MM may only be used for interpolations. That is, force fields that are entirely based on structural data may not lead to reasonably accurate predictions of conformational equilibria. However, in general force fields, where parts of the parameterization schemes have been fitted to thermodynamic data, it has been found that reasonable agreement of relative strain energies with experimentally determined stabilities is observed, and this is not unexpected [13,20–24].

5.3. The environment

When structures are optimized by MM the computation very often does not involve interactions with neighboring molecules. However, when the parameterization of the entire force field is based on solid state (or solution) data the environment is implicitly included in the model. That is, the computed molecular structures and properties are not from 'gas phase' molecules (sometimes they are called "interactionless", and this may be more appropiate), and the environmental effects are isotropic [20].

6. Interpretation and communication

The limits of the MM approach have been stressed above. This is the basis for the interpretation of the results. The main factor is that MM is a purely interpolative approach. It follows that full details of the force field have to be presented together with the results and their interpretation, in order to allow the reader to fully reproduce them [43]. These include (i) the functional forms of the potentials; (ii) the parameter sets used; (iii) the data for which they were derived (e.g. set of X-ray structures from the CSD) and (iv) the data for which they were validated.

Acknowledgements

We are grateful to the Alexander von Humboldt Foundation for generous financial support.

References

[1] S.R. Niketic, K. Rasmussen, The Consistent Force Field, vol. 3, Springer, Germany, 1977.
[2] U. Burkert, N.L. Allinger, Molecular Mechanics, vol. 177, American Chemical Society, New York, 1982.
[3] G.R. Brubaker, D.W. Johnson, Coord. Chem. Rev. 53 (1984) 1.
[4] J.C.A. Boeyens, Struct. Bonding 63 (1985) 65.
[5] R.D. Hancock, Progr. Inorg. Chem. 37 (1989) 187.
[6] T.W. Hambley, Commun. Inorg. Chem. 14 (1992) 1.
[7] P. Comba, Coord. Chem. Rev. 123 (1993) 1.
[8] B.P. Hay, Coord. Chem. Rev. 126 (1993) 177.
[9] P. Comba, Commun. Inorg. Chem. 16 (1994) 133.
[10] R.J. Deeth, Struct. Bonding (Berlin) 82 (1995) 1.
[11] C.R. Landis, D.M. Root, T. Cleveland, in: K.B. Lipkowitz, D.B. Boyd (Eds.), Reviews in Computational Chemistry, VCH, Weinheim, 1995, p. 73.
[12] M. Zimmer, Chem. Rev. 8 (1995) 2629.
[13] P. Comba, T.W. Hambley, Molecular Modeling of Inorganic Compounds, VCH, Weinheim, 1995; second ed., Wiley-VCH, Weinheim, 2000.
[14] K.B. Lipkowitz, J. Chem. Educ. 72 (1995) 1070.
[15] W. Gans, A. Amann, J.C.A. Boeyens, Fundamental Principles of Molecular Modeling, Plenum Press, New York, 1996.
[16] P. Comba, M. Zimmer, J. Chem. Educ. 73 (1996) 108.
[17] A.K. Rappé, C.J. Casewit, Molecular Mechanics across Chemistry University Science Books, Sausalito, 1997.
[18] L. Banci, P.E. Comba, Molecular Modeling and Dynamics of Bioinorganic Systems, Kluwer, Dordrecht, 1997.
[19] T.R. Cundari, J. Chem. Soc. Dalton Trans. (1998) 2771.
[20] P. Comba, in: W. Gans, J.C.A. Boeyens (Eds.), Intermolecular Interactions, Plenum Press, New York, 1998.
[21] P. Comba, Coord. Chem. Rev. 182 (1999) 343.
[22] P. Comba, Coord. Chem. Rev. 185 (1999) 81.
[23] P. Comba, in: J.K.A. Howard, F.H. Allen (Eds.), Implications of Molecular and Materials Structure for New Technologies, Kluwer, Dordrecht, 1999, p. 71.

[24] P. Comba, in: J.K.A. Howard, F.H. Allen (Eds.), Implications of Molecular and Materials Structure for New Technologies, Kluwer, Dordrecht, 1999, p. 87.
[25] J.C.A. Boeyens, F.M.M. O'Neill, Inorg. Chem. 37 (1998) 5346.
[26] J.C.A. Boeyens, Inorg. Chem. 24 (1985) 4149.
[27] P. Comba, in: W. Gans, A. Amann, J.C.A. Boeyens (Eds.), Fundamental principles of molecular modeling, Plenum Press, New York, 1996, p. 167.
[28] A. Julg, Croatica Chim. Acta 57 (1984) 1497.
[29] J.C.A. Boeyens, in: W. Gans, A. Amann, J.C.A. Boeyens (Eds.), Fundamental priciples of molecular modeling, Plenum Press, New York, 1996, p. 167.
[30] J.E. Bol, C. Buning, P. Comba, J. Reedijk, M. Ströhle, J. Comput. Chem. 19 (1998) 512.
[31] P. Comba, T.W. Hambley, M. Ströhle, Helv. Chim. Acta 78 (1995) 2042.
[32] R.J. Deeth, I.J. Munslow, V.J. Paget, in: L. Banci, P. Comba (Eds.), Molecular Modeling and Dynamics of Bioinorganic Systems, Kluwer, Dordrecht, 1997.
[33] C.R. Landis, T.K. Firman, T. Cleveland, D.M. Root, in: L. Banci, P. Comba (Eds.), Molecular Modeling and Dynamics of Bioinorganic Systems, Kluwer, Dordrecht, 1997, p. 49.
[34] P. Comba, M. Zimmer, Inorg. Chem. 33 (1994) 5368.
[35] S.L. Mayo, B.D. Olafson, A. Goddard III, J. Phys. Chem. 94 (1990) 8897.
[36] A.K. Rappé, K.S. Colwell, C.J. Casewit, Inorg. Chem. 32 (1993) 3438.
[37] A.S. Lopis, L. Glasser, F. Marsicano, QCPE Bull. 17 (1997) 16.
[38] J.L.M. Dillen, J. Comput. Chem. 13 (1992) 257.
[39] P.O. Norrby, T. Liliefors, J. Comput. Chem. 19 (1998) 1146.
[40] J. Hunger, S. Beyreuther, G. Huttner, K. Allinger, U. Rachlow, L. Zsolnai, Eur. J. Inorg. Chem. (1998) 693.
[41] J. Bartol, P. Comba, M. Melter, M. Zimmer, J. Comput. Chem. 20 (1999) 1549.
[42] J.C.A. Boeyens, Acta Crystallogr. Sect. C 39 (1983) 846.
[43] D.J. Raber, W.C. Guida, J. Comput. Chem. 20 (1998) 1591.