Empirical Molecular-Hydrogen Wave Function Modeled from Theoretically Derived X-Ray Diffraction Data

R. F. Boehme and S. J. La Placa
IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598

(Received 28 April 1987)

X-ray empirical wave-function models (XEWM's) for molecular hydrogen have been obtained by use of a density-matrix formalism fitting to x-ray diffraction data. Data are derived from the Kolos-Roothaan wave function and an imposed crystal lattice. Molecular properties calculated from our XEWM's are in excellent agreement with theoretical results. An XEWM with a crystallographic residual of $R_{wF}=0.0007$ yields a calculated binding energy of 0.1332 a.u., extremely close to the Hartree-Fock limit of 0.1336 a.u. This study demonstrates the accuracy with which XEWM's are able to predict molecular properties from x-ray diffraction data.

PACS numbers: 31.20.Pv, 61.10.−i

A high-accuracy x-ray empirical wave-function model (XEWM) for molecular hydrogen has been extracted at the Hartree-Fock level from x-ray diffraction data. This represents a substantial and fundamental step in the ability of the x-ray crystallographers to predict properties and to correlate structure and bonding. Properties calculated from our XEWM's include the electric field, the electric field gradient, the quadrupole moment, and the total energy for molecular hydrogen. The quality of the above properties depends on the flexibility of the parametrized model wave function chosen (e.g., basis functions, symmetry restrictions, and allowed orbital overlaps). Therefore they critically test our ability to extract meaningful and accurate XEWM's. The limited amount of data available from an x-ray diffraction experiment demands that XEWM's have relatively few parameters compared with theoretical methods. Fortunately, the linear combination of atomic orbitals (LCAO) approach provides a sufficiently simple description that an accurate XEWM can be constructed, and it also provides a model which can be easily analyzed. Further, the x-ray diffraction data provide an important restriction, the Fourier transform of the total electron density, not available to purely theoretical methods.

The conventional model for x-ray structure factors takes the form of x rays being coherently scattered from independent nonbonded spherical atoms. The parameters of atomic types, positions, and thermal motions are determined by a least-squares fit to the observed x-ray structure factors, $F_0(h)$. The ability of the crystallographers to obtain high-precision $F_0(h)$, the Fourier transform of which clearly shows the electron distribution due to bonding, has stimulated great interest in extracting high-precision information about the properties and bonding of crystal components. The most widely accepted models for analysis of high-precision $F_0(h)$ have been multipole pseudoatom models. Whereas excellent fits to $F_0(h)$, and therefore electron density, are achieved, there are severe limits to the analysis of the results inherent in these models.

This work is based on the premise that a wave-function model is best to predict molecular properties, to extract molecules from crystalline solids, and to describe high-precision $F_0(h)$. It is at this level where bonding can be directly described, all electronic properties can be predicted, and the constraints which assure a physically meaningful model can be applied.

Stewart's[1] development of a generalized x-ray scattering formalism for improved fits to $F_0(h)$ provides the necessary framework. The one-electron density $\rho(r)$ and therefore the coherent x-ray scattering can be ascribed to arise from molecular wave functions Φ_i which are formed from linear combinations of atomic orbitals Ψ_j with coefficients C_{ij}:

$$\Phi_i=\sum_j C_{ij}\Psi_j \quad \text{or} \quad \mathbf{\Phi}=\mathbf{C}\mathbf{\Psi} .$$

$\rho(r)$ is given by

$$\rho(r)=2\sum_i \Phi_i^*(r)\Phi_i(r)=2\sum_j\sum_k P_{jk}\Psi_j^*(r)\Psi_k(r)$$
$$=2\,\mathrm{Tr}\mathbf{P}\mathbf{\Psi}(r)\mathbf{\Psi}^\dagger(r) ,$$

where $P_{jk}=\sum_i C_{ij}C_{ik}$ or the population matrix $\mathbf{P}=\mathbf{C}^\dagger\mathbf{C}$. Similarly the calculated model x-ray structure factors are $F_C(h)=2\,\mathrm{Tr}\mathbf{P}\mathbf{f}(h)$, where $\mathbf{f}(h)$ is a matrix whose elements are x-ray scattering products due to all possible atomic orbital products for each diffraction index h.

Coppens and co-workers[2-5] applied Stewart's formalism by least-squares refinement of selected elements of $\mathbf{P}$. Problems arose because of the number of parameters and their correlation both within the population matrix and with conventional crystallographic parameters.

Importantly, it was demonstrated by Massa and co-workers[6-8] that proper fundamental constraints on the population matrix are necessary to achieve physically meaningful model wave functions. These constraints reduce the number and correlation of parameters within the population matrix. Additionally, Massa and co-workers showed that pseudoatom multipole models do not satisfy the Pauli principle nor necessarily result in total electron densities which are everywhere nonnegative, constraints which are fundamental to any electronic wave-function model. The matrix projector constraint $\mathbf{P}^2=\mathbf{P}$ assures that the wave functions are orthonormal, and the condition $\mathrm{Tr}\mathbf{P}=N$ defines the number of occupied wave functions equal to N for orthonormal basis functions.

Previous work on beryllium[9] demonstrates that excellent fits to experimentally observed structure factors and electron densities can be achieved with use of a model empirical wave function restricted to rehybridization of orbitals on a single atomic site describing independent nonbonded atoms in a crystal field. This model successfully deconvolutes conventional crystallographic parameters of scale factor and thermal motion from the quantum-based population matrix parameters.

This Letter demonstrates for a molecular crystal that not only excellent density fits and deconvolution of parameters can be achieved, but in addition that a molecular density can be extracted from a crystalline density and that an empirical wave-function model based on x-ray diffraction data can accurately predict bonding and properties for this isolated molecule.

In our present work all the elements of $\mathbf{P}$ are determined by matrix solution of linear equations related to the $F_0(h)$ and the necessary population matrix constraints.[10] In a single iterative step each new population matrix element P'_{ij} is defined as the sum of the initial element P_{ij} plus the shift, Δ_{ij}, necessary to improve the fit: $P'_{ij}=P_{ij}+\Delta_{ij}$. The equations, in explicit form, necessary to solve for the Δ_{ij} are as follows: (1) one equation for each observation, $F_0(h)$:

$$w(h)[2\textstyle\sum_{ij}\Delta_{ij}f_{ji}(h)] = w(h)[F_0(h)-2\textstyle\sum_{ij}P_{ij}f_{ji}(h)] \; ,$$

where $w(h)$ are weights related to the reliability of each $F_0(h)$; (2) one equation for each element P_{ij} to enforce orthonormality on the model wave functions:

$$w_c[\Delta_{ij}-\textstyle\sum_k(P_{ik}\Delta_{kj}+\Delta_{ij}P_{kj})] = w_c(-P_{ij}+\textstyle\sum_k P_{ik}P_{kj}) \; ,$$

where w_c is the weight necessary to enforce constraints; (3) the number of occupied functions is determined by $w_c(\sum_k\Delta_{kk})=w_c(N-\sum_k P_{kk})$. The conventional crystallographic parameters of atomic position (and therefore internuclear distance), thermal motion, and crystallographic scale factors are determined by conventional crystallographic least squares.

The x-ray diffraction data are calculated *a priori*, based on the correlated wave function of Kolos and Roothaan (KR).[11] Davidson and Jones (DJ)[12] extracted the first ten natural orbitals of the KR wave function, which were then used by Stewart, Davidson, and Simpson (SDS)[13] to calculate x-ray scattering data for the hydrogen molecule. The SDS-tabulated entries were used to form partial x-ray diffraction data sets with a $\sin\theta/\lambda$ maximum of 1.5028 Å^{-1} based on equiaxial orthogonal unit cells with $a=23.290$, 11.645, or 5.823 Å and one hydrogen molecule per unit cell centered about the origin with atomic positions at $\pm z$. As there is no significant dependence of XEWM's on unit-cell size, we report only those for the 23-Å unit cell with 279 independent diffraction intensities. The data used in these refinements are on an absolute scale and contain no errors due to absorption or extinction of x rays, nor do they contain any thermal motion effects.

Table I contains the LCAO coefficient, C_{ij}, and exponential scaling factors, ζ, for four different XEWM solutions, each varying in the particular basis function available to produce optimal fits to the $F_0(h)$. The basis function consists of normalized sums of normalized Gaussians. All refinements utilize Huzinaga's[14] atomic-hydrogen 1s Slater-type orbital which is a sum of ten Gaussian functions (STO-10G):

$$\Phi_{1s}=N\sum_{i=1}^{10}a_i(2\zeta^2\alpha_i/\pi)^{3/4}\exp(-\zeta^2\alpha_i r^2) \; ,$$

where the a_i and α_i are from Huzinaga. The ζ are fixed at 1 in refinements for models 1 and 2 and are optimized in refinements for models 3 and 4 to allow for radial scaling of the basis functions. Models 2 and 4 provide additional flexibility by splitting Huzinaga's STO-10G into two functions of 1s symmetry (each with an independent ζ) consisting of the seven sharpest (7G) and

TABLE I. LCAO coefficients C, exponential scaling factors ζ, and distance of $2p_z$ Gaussian lobes from nuclear position r, for atomic basis functions obtained in XEWM solutions of σ_g symmetry.

	1s-10G		1s-7G		1s-3G		2p-3G		2p-1G		r
Model	C	ζ	C	ζ	C	ζ	C	ζ	C	ζ	(Å)
1	0.5049	1[a]	···	···	···	···	···	···	···	···	···
2	···	···	0.2225	1[a]	0.3600	1[a]	···	···	···	···	···
3	0.5398	1.188	···	···	···	···	···	···	···	···	···
4	···	···	0.3030	1.122	0.2825	1.157	0.0313	4.656	0.0403	0.442	0.154

[a] Fixed in refinement.

TABLE II. Properties calculated for XEWM solutions and for Davidson-Jones (DJ) wave function. As the crystallographic R_{wF} improves, the XEWM properties converge to accurately predict the DJ properties.

Model	R_{wF} (%)	D (Å)	Bond energy (a.u.)	Electric field (a.u.)	Electric field gradient (a.u.)	Quadrupole moment (a.u.)
1	9.52	0.259	−0.8057	3.185	−16.64	0.05
2	0.66	0.686	0.1255	0.132	−0.55	0.28
3	0.57	0.683	0.1252	0.141	−0.56	0.28
4	0.07	0.736	0.1332	0.012	−0.37	0.43
DJ	0	0.741	0.1336[a]	0.002	−0.34	0.46

[a] Hartree-Fock limit rather than DJ value of 0.1744.

three most diffuse (3G) Gaussians. Polarization is added in model 4 by the addition of two Gaussian-lobe P_z basis functions. The first is Whitten's[15] $2P_z$ for atomic hydrogen with three Gaussians per lobe. The second consists of a single Gaussian per lobe. Both P_z functions have the Gaussian lobes at a single distance along the bond axis from the center of the $1s$ functions with their positive lobes in the interatomic region.

Results in Table II show that physical properties for molecular hydrogen can be accurately predicted based on the XEWM's extracted by fitting to $F_0(h)$'s for a crystal containing hydrogen molecules. Each improvement in the set of basis functions may be seen equally well by the improvement in the crystallographic R factor,

$$R_{wF}=\sum_h w(h)\left|\,|F_0(h)|-|F_C(h)|\,\right|^2/\sum_h w(h)|F_0(h)|^2$$

or by the bond energy calculated with a Hartree-Fock Hamiltonian. Indeed XEWM's predict bond energies very close to those obtainable with the same basis functions in an energy minimization. Model 4 is at the Hartree-Fock limit for the bond energy, 0.1332 a.u. vs 0.1336 a.u. From a purely crystallographic point of view, each improvement in the model and R_{wF} yields substantial improvement in the predicted H-H bond distance, D. This method therefore promises improved atomic coordinates and bond distances from crystallographic studies, especially for atoms with "lone pairs" and for bonds to hydrogen atoms.[16]

Analysis of high-quality x-ray diffraction data has included calculation of electrostatic properties in the valence regions of molecules. Therefore we have calculated the electric field, the electric field gradient, and quadrupole moment for molecular hydrogen at the nucleus. Again XEWM's make predictions appropriate for the quality of the model, with models 2 and 3 providing good results and model 4 providing excellent results.

This study has shown that an x-ray empirical wave-function model can be extracted from coherent x-ray diffraction data for molecular hydrogen, and that this model can be used to accurately predict molecular properties for this system.

---

[1] Robert F. Stewart, J. Am. Chem. Soc. **51**, 4569–4577 (1969).

[2] P. Coppens, T. V. Willoughby, and L. N. Csonka, Acta Crystallogr., Sect. A **27**, 248–256 (1971).

[3] P. Coppens, D. Pautler, and J. F. Griffin, J. Am. Chem. Soc. **93**, 1051–1058 (1971).

[4] D. A. Matthews, G. D. Stucky, and P. Coppens, J. Am. Chem. Soc. **94**, 8001–8008 (1972).

[5] D. S. Jones, D. Pautler, and P. Coppens, Acta Crystallogr., Sect. A **28**, 635–645 (1972).

[6] William L. Clinton and Louis J. Massa, Phys. Rev. Lett. **29**, 1363–1366 (1972).

[7] C. A. Frishberg and Louis J. Massa, Phys. Rev. B **24**, 7018–7024 (1981).

[8] C. A. Frishberg and Louis J. Massa, Acta Crystallogr., Sect. A **38**, 93–98 (1982).

[9] L. J. Massa, M. J. Goldberg, C. Frishberg, R. F. Boehme, and S. J. La Placa, Phys. Rev. Lett. **55**, 622–625 (1985).

[10] R. F. Boehme, to be published.

[11] W. Kolos and C. C. J. Roothaan, Rev. Mod. Phys. **32**, 219–232 (1960).

[12] Ernest R. Davidson and Leon L. Jones, J. Chem. Phys. **37**, 2966–2971 (1962).

[13] Robert F. Stewart, Ernest R. Davidson, and William T. Simpson, J. Am. Chem. Soc. **42**, 3175–3187 (1965).

[14] Sigeru Huzinaga, J. Chem. Phys. **42**, 1293–1302 (1965).

[15] Jerry L. Whitten, J. Chem. Phys. **39**, 349–362 (1963).

[16] F. H. Allen, Acta Crystallogr., Sect. B **42**, 515–522 (1986).