JOURNAL OF PHYSICAL ORGANIC CHEMISTRY
J. Phys. Org. Chem. 2003; **16:** 133–141
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/poc.584

Description of covalent bond orders using the charge density topology

Siân T. Howard* and Olivier Lamarche

Department of Chemistry, Cardiff University, Cardiff CF10 3TB, UK

Received 9 July 2002; revised 10 September 2002; accepted 16 September 2002

epoc

ABSTRACT: Hybrid density functional calculations are employed to explore the relationships between covalent bond order, as defined using the atomic overlap matrix (AOM) formalisms developed by Cioslowski, Ángyán and others, and the parameters derived from a topological analysis of the electron density. Relationships are obtained for the specific cases of C—C, C—N, C—O, C—P and C—S bonds. The simple Pauling bond order–bond length relationship describes the data reasonably well in most cases, but the correlations show considerable scatter. Although no single parameter acts as a unified descriptor of bond order for all types of bond, in each case it is possible to find a model which describes the bond order data significantly better than the Pauling model based on bond length. The relationships presented can therefore be utilized to estimate rapidly the covalent part of the bond order from a topological analysis of the charge distribution for very large systems where the AOM-based methods can become impractical to apply, and for charge density distributions which have been obtained from experiment (e.g. elastic x-ray scattering). Copyright © 2003 John Wiley & Sons, Ltd.
Additional material for this paper is available from the epoc website at http://www.wiley.com/epoc

KEYWORDS: covalent bond orders; charge density topology

INTRODUCTION

The concept of bond order has played a key role in the development of chemical bonding theories.[1] Bond orders have since been applied more widely, for example in the study of chemical reactions[2] (including bond order conservation),[3,4] as a tool in inorganic chemistry[5] and in developing simple models of bond energies[6] and bond entropies.[7] The earliest and most widely known attempt by quantum chemists to provide a quantitative measure of bond order is Pauling's two-parameter description in terms of bond length:[1]

$$n = \exp[(r_0 - r)/a] \tag{1}$$

where a ≈0.3 for essentially any type of bond, and r_0 is an idealized single bond length for the type of bond in question. As an illustration, we take theoretical data for the optimized C—C bond lengths in ethane, ethene and ethyne (B3LYP/6–31 + G** level) of 1.532, 1.334 and 1.208 å, respectively, and assign bond orders of 1, 2 and 3 to the C—C bonds in these compounds, which readily gives [see Fig. 1(a)]

$$n(r) = \exp[(1.538 - r)/0.299] \tag{2}$$

Thus a bond order for any C—C bond optimized at the same level of theory can be estimated by simply inserting its bond length into Eqn. (1). Lendvay[8] used minimal basis Hartree–Fock calculations to derive values for the exponential dependence in Pauling's relationship, and found a(C—C bonds) ≈ 0.37, a(C—O) ≈ 0.38 and a(C—H) ≈ 0.26.

Bader[9] argued, and most quantum chemists would accept, that a bond order index should properly be based on the electron distribution rather than the internuclear distance (bond length), since it is the electrons that are actually doing the bonding. (Indices based on bond length presumably work reasonably well because of the strong correlation with many density-derived parameters and bond length.) By analogy with Pauling's relationship, Bader proposed a bond order index based on the bond critical point (CP) electron density (the value of the electron density at the point in the bond where $\nabla\rho = 0$):

$$n(\rho_c) = \exp[a(\rho_c - b)] \tag{3}$$

Using the B3LYP/6–31 + G** charge density topological data presented in Table 1 for ethane, ethene and ethyne, we can again illustrate this relationship [see Fig.

**Correspondence to:* S. T. Howard, Department of Chemistry, Cardiff University, Cardiff CF10 3TB, UK.
E-mail: howardst@cf.ac.uk

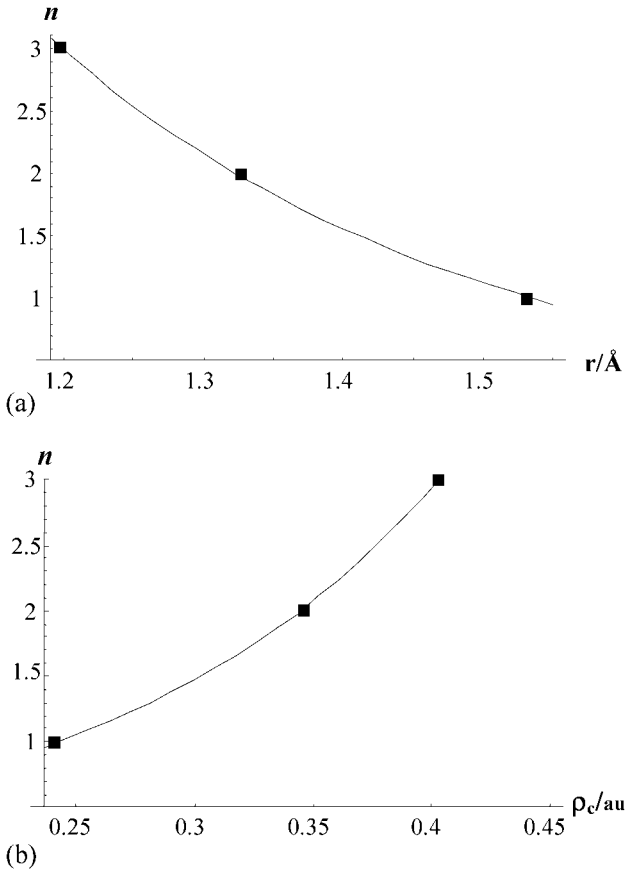

Figure 1. Illustration of (a) the Pauling and (b) the Bader bond-order relationships using B3LYP/6–31 + G** data for ethane, ethene and ethyne

1(b)] by assigning bond orders of 1, 2 and 3 to the C—C bonds of these compounds, giving

$$n(\rho_c) = \exp[6.862(\rho_c - 0.243)] \tag{4}$$

Clearly, the extent to which the predictions of Eqn. (2) based on internuclear distance and Eqn. (4) based on cricical point density agree will depend on the how closely the critical point density and bond length are correlated for a given type of bond.

In the context of charge density topological properties, the work of Silvi and Savin[10] using the electron localization function (ELF) should be mentioned. Although their work suggested a link between bond multiplicity and the ELF, it remains to be shown that a practical, quantitative bond order index can be based on this function.

Practical definitions of bond order based on results of quantum mechanical calculations date back to Wiberg's index, which applies to semi-empirical methods with unit overlap matrices.[11] Briefly, given a set of LCAO canonical orbitals $\{\phi_i = \sum_j c_{ij}\ \chi_j$ which are eigenfunctions of SCF Fock or Kohn–Sham operators ($f_i\ \phi_i = \epsilon_i\ \phi_i$), the electron density expression is

$$\rho(\mathbf{r}) = \sum_{k=1}^{N_{occ}} n_k \phi_k(\mathbf{r})^* \phi_k(\mathbf{r}) = \sum_k n_k \sum_i \sum_j c_{ij} c_{ik} \chi_i(\mathbf{r}) \chi_j(\mathbf{r})$$
$$= \sum_i \sum_j P_{ij} \chi_i(\mathbf{r}) \chi_j(\mathbf{r}) \tag{5}$$

where n_k is the occupation number of the kth occupied molecular orbital. Integration of Eqn. (5) over $\mathbf{r}$ gives the total number of electrons in the system, $N = PS$. The Wiberg index n_W assumes that the AO basis function overlap matrix S is the unit matrix, and relates the bond order to a sum over squared density matrix elements linking pairs of basis functions on the two atomic centres (A and B):

$$n_W = \sum_{i \in A} \sum_{j \in B} P_{ij}^2 \tag{6}$$

Although it has occasionally been used to calculate bond orders using *ab initio* wavefunction with explicit (non-unit) overlap matrices, perhaps because its ease of computation, it is only really suited to ZDO semiempirical calculations. Mayer[12] subsequently presented a bond order index which has become very popular, and takes proper account of the AO overlap. In the closed-shell case it is given by

$$n_M = \sum_{i \in A} \sum_{j \in B} (P_{ij} S_{ij})(P_{ji} S_{ji}) \tag{7}$$

More recently, Cioslowski and Mixon[13] proposed a bond order index which has its basis in the atoms-in-molecules method developed by Bader.[9] The partitioning scheme produces atoms (strictly atomic basins Ω) enclosed by atomic surfaces with normal vectors a such that $\nabla\rho \cdot \mathrm{a} = 0$: these atoms obey the Virial theorem $2T(\Omega) = V(\Omega)$. An 'atomic overlap matrix' (AOM)[14] element is defined in the MO basis as the integral of a

Table 1. Illustrative data for C—C bonds (B3LYP/6–31 + G** level)

	n_C[a]	ρ_c (e bohr^{-3})	λ_1 (e bohr^{-3})	λ_2 (e bohr^{-3})	λ_3 (e bohr^{-3})	$\nabla^2\rho_c$ (e bohr^{-3})
$H_3C—CH_3$	1.032	0.241	−0.444	−0.444	0.400	−0.548
$H_2C=CH_2$	1.936	0.346	−0.745	−0.548	0.277	−1.016
HC≡CH	2.895	0.403	−0.599	−0.599	0.027	−1.171

[a] The Cioslowski–Mixon bond order.

basis function product over a particular basin (A):

$$< \phi_i|\phi_j >_{\mathrm{A}} = \int_{\Omega} \phi_i(\mathbf{r})\phi_j(\mathbf{r})\mathrm{d}\mathbf{r} \qquad (8)$$

Cioslowski and Mixon[13] argued that bond orders would be best computed in a localized orbital basis which maximizes the sum of the squared diagonal elements of each AOM. Having obtained these localized orbitals, the bond order in the closed-shell case is computed from

$$n_{\mathrm{C}} = \sum_k {n_k}^2 < \phi_k|\phi_k >_{\mathrm{A}} < \phi_k|\phi_k >_{\mathrm{B}} \qquad (9)$$

Table 1 reports B3LYP/6–31 + G** values of n_{C} for the ethane, ethene and ethyne series. The values are all fairly close to the nominal values of 1, 2 and 3 for these (non-polar) bonds. However, since n_{C} reflects the (dominant) *covalent* part of the bond order, bonds with substantial ionic character can show values of these indices much lower than the nominal integers. Ángyán *et al.*[15] subsequently derived a covalent bond order based on AOMs which does not depend on any particular choice of localized orbitals:

$$n_{\mathrm{A}} = \sum_i \sum_j n_i n_j < \phi_i|\phi_j >_{\mathrm{A}} < \phi_j|\phi_i >_{\mathrm{B}} \qquad (10)$$

In many cases, these two indices give similar values, but they can differ considerably for aromatic molecules.[15] It should also be re-emphasized that both n_{C} and n_{A} omit the ionic contributions to the total bond order, so they should not be used (for example) to measure total bond strength. In this work, we use the Cioslowski–Mixon formalism to generate a body of covalent bond order data to be modelled with bond CP topological properties. The latter have been applied as descriptors of a wide range of molecular properties, including electronegativity,[16] bond energy,[17,18] aromaticity,[19] hydrogen bond strength,[20] hydrogen bond donor capacity[21] and molecular similarity.[22]

The main aim of this work was to explore in more depth Bader's proposal of using bond charge density topological properties such as ρ_{c} as descriptors of bond order. In doing so, we will also provide the most thorough test of the Pauling bond order–bond length relationship so far reported. It will be shown that, provided other bond charge density descriptors are considered such as $\nabla^2\rho_{\mathrm{c}}$ and the eigenvalues of the Hessian of ρ, $\{\lambda_1, \lambda_2, \lambda_3\}$, then significantly better correlations with covalent bond order are found than those offered by the basic expressions due to Pauling and Bader. Simple models of bond order utilizing the kinetic energy density at the bond critical point, viz.

$$G_{\mathrm{c}} = \frac{1}{2} \sum_{k=1}^{N_{\mathrm{occ}}} (\nabla\phi_k) \cdot (\nabla\phi_k) \qquad (11)$$

will also be considered. A secondary but important aim is to provide a tool for estimating bond order for densities where no wavefunction is available, such as the many-centre multipole-parameterized charge densities which can be obtained from high-resolution elastic x-ray scattering experiments.[23]

In any multivariate statistical description of a data set, it is important to take into account the inter-correlations between descriptors, and these will be considered at each stage of generating various models of the bond order data. This in turn generates useful information concerning any future attempt to model a property using bond CP data, since relatively little is known about the extent to which these parameters are independent of one another.

COMPUTATIONAL METHOD

We considered bonds between carbon and five elements with which it forms multiple bonds: C—C, C—N, C—O, C—S and C—P. A comprehensive set of species were chosen for each type of bond (containing 37 bonds of each type) in order to represent a wide range of 'embedding environments.' In each case, this set included a number of cations, anions and both neutral and charged radicals as well as the more common neutral closed-shell species. All species were geometry optimized and frequency tested using the B3LYP hybrid functional[24,25] and 6–31 + G** basis sets[26] as employed in Gaussian 98.[27] Densities were analysed at this same level of theory. The bond orders and topological analyses of all compounds were carried out using the atoms-in-molecules algorithms due to Cioslowski and co-workers implemented in the Gaussian code. The key data extracted from these analyses are included in the Electronic Supplementary Information at the epoc website at http://www.wiley.com/epoc; more complete details of the compounds and density analyses may be obtained from the author (S.T.H.) on request.

In addition to the bond critical point electron density suggested by Bader, we focus on $\nabla^2\rho_{\mathrm{c}}$ and in particular its constituent principal curvatures $\nabla^2\rho_{\mathrm{c}} = \partial^2\rho_{\mathrm{c}}/\partial x^2 + \partial^2\rho_{\mathrm{c}}/\partial y^2 + \partial^2\rho_{\mathrm{c}}/\partial z^2 = \lambda_1 + \lambda_2 + \lambda_3$ as descriptors of bond order. The rationale for this is that ρ_{c} only indirectly contains information about π-bonding (although the bond CP is on or close to the nodal plane for π orbitals, σ and π electron density distributions are necessarily coupled via the self-consistent field procedure). The principal curvatures, on the other hand, are anisotropic measures of electron density concentration/depletion perpendicular to the bond (λ_1 and λ_2) and along the bonding direction (λ_3). For this reason they can be expected to be more sensitive to details of π and σ electron distributions, respectively. This is suggested by the charge density topology of benzenoid aromatic compounds[19] which show maximum correlation of aromaticity indices with the curvature perpendicular to the planar ring systems. It is therefore

Table 2. Summary of bond order models for C—N, C—O and C—P bonds

	C—N bonds				C—O bonds				C—P bonds			
Model 1, $n = a + b\nabla^2\rho_c + c\rho_c$:												
Coefficients	a	b	c		a	b	c		a	b	c	
	−0.175	0.376	5.409		0.569	0.276	1.847		1.042	1.397	0.0	
Statistics	R^2	R^2_{CV}	r.m.s.	*F*-stat	R^2	R^2_{CV}	r.m.s.	*F*-stat	R^2	R^2_{CV}	r.m.s.	*F*-stat
	0.975	0.969	0.099	657.0	0.926	0.898	0.086	214.1	0.918	0.910	0.188	393.3
Model 2, $n = a + b\lambda_3 + c\lambda_{12} + d\rho_c$:												
Coefficients	a	b	c	d	a	b	c	d	a	b	c	d
	−0.284	0.331	0.559	6.569	0.776	0.267	0.0	0.0	−0.107	1.022	5.136	17.30
Statistics	R^2	R^2_{CV}	r.m.s.	*F*-stat	R^2	R^2_{CV}	r.m.s.	*F*-stat	R^2	R^2_{CV}	r.m.s.	*F*-stat
	0.977	0.968	0.096	467.1	0.932	0.922	0.081	480.0	0.939	0.918	0.167	169.4
Model 3, $n = a + b\nabla^2\rho_c + c\rho_c + dX$:												
Coefficients	a	b	c	$d\rho^2$	a	b	c	$d\rho^2$	a	b	c	$d(\nabla^2\rho)^2$
	−1.401	0.456	13.20	−11.39	−0.626	0.482	10.60	−14.33	0.993	0.924	0.00	0.661
Statistics	R^2	R^2_{CV}	r.m.s.	*F*-stat	R^2	R^2_{CV}	r.m.s.	*F*-stat	R^2	R^2_{CV}	r.m.s.	*F*-stat
	0.978	0.968	0.094	495.6	0.952	0.937	0.070	217.2	0.942	0.933	0.160	277.2
Model 4, $n = a + b\ G_c$:												
Coefficients	a	b			a	b			a	b		
	0.745	1.861			0.644	1.002			0.286	5.222		
Statistics	R^2	R^2_{CV}	r.m.s.	*F*-stat	R^2	R^2_{CV}	r.m.s.	*F*-stat	R^2	R^2_{CV}	r.m.s.	*F*-stat
	0.972	0.969	0.103	1215.2	0.910	0.899	0.093	355.3	0.882	0.871	0.226	262.1
Pauling, $n = \exp[(r_0 - r)/a]$:												
Coefficients	r_0	a			r_0	a			r_0	a		
	1.435	0.308			1.353	0.410			1.810	0.290		
Statistics	R^2	r.m.s.	*F*-stat		R^2	r.m.s.	*F*-stat		R^2	r.m.s.	*F*-stat	
	0.961	0.123	851.7		0.857	0.118	210.2		0.876	0.231	247.9	
Bader, $n = \exp[(\rho_c - b)/a]$:												
Coefficients	b	a			b	a						
	0.277	0.207			0.302	0.319						
Statistics	R^2	r.m.s.	*F*-stat		R^2	r.m.s.	*F*-stat					
	0.914	0.181	373.5		0.842	0.124	186.8					

possible that $(\lambda_1 + \lambda_2)$ and λ_3 may be sensitive to the separate π and σ contributions to the bond order, respectively. Note that in the case of the two perpendicular curvatures we use the sum value $(\lambda_1 + \lambda_2)$, which for brevity will be denoted λ_{12}.

In principle, the bond ellipticity $\epsilon = \lambda_1/\lambda_2 - 1$ could also be used as a descriptor of covalent bond order. However, ε has the well-known behaviour of increasing to a maximum value for a 'perfect' double bond, and then decreasing to zero in the range of double–triple bonds as the bond approaches cylindrical C_∞ symmetry. In other words, the bond order is not a single-valued function of ε, so ε is not suitable as a statistical descriptor of bond orders between atoms which can form triple bonds, e.g. C—C, C—N, C—P and so on. We want to construct a model with general applicability, so ε has not been used as a descriptor here.

Multiple (stepwise) linear and non-linear regression was used to model the bond order data using JMP[28] and Mathematica 4,[29] with a variety of descriptors. The quality of each model is judged using four statistics: R^2 (correlation coefficient), equal to the fraction of variance accounted for by the model; cross-validated (or 'leave-one-out') correlation coefficient R^2_{CV}; the usual root-mean-square error, r.m.s.; and Fischer's *F*-statistic, *F*-stat, which measures the significance of the model given the number of variables employed.

RESULTS

The observed range of B3LYP/6–31 + G** covalent bond orders (full detail are given in the Electronic Supplementary Information) is as follows: 0.93 (in $H_3C—CO_2^-$) $<n_C$ (C—C) <2.90 (in HC≡CH); 0.90 (in $[H_2C—NHCH_3]^+$) $<n_C$ (C—N) <2.67 (in C≡N·); 0.76 (in $[H_3C—OH_2]^+$) $<n_C$ <(C—O) <1.83 (in $[P{=}C{=}O]^+$); 0.72 (in $[H_3P—C{\equiv}N]^+$) $<n_C$ (C—P) <2.52 (in P≡CH); and 0.98 (in Me_2SO_2) $<n_C$ (C—S) <2.78 (in $[HC{=}S\cdot]^+$). The phosphorus compounds included examples of P—C bonds to phosphorus(V) and phosphonium ylides; the sulphur compounds contained examples of S—C bonds to sulphur(VI). Cyclic and cage compounds were also included where possible. In general, every effort was made to ensure that a range of 'exotic' bond types was included in each data set, to

Table 3. Summary of bond order models for C—C and C—S bonds

	C—C bonds				C—S bonds			
Model 1, $n = a + b\nabla^2\rho_c + c\rho_c$:								
Coefficients	a	b	c		a	b	c	
	−1.865	0	11.34		−0.460	0.684	10.17	
Statistics	R^2	R^2_{CV}	r.m.s.	F-stat	R^2	R^2_{CV}	r.m.s.	F-stat
	0.945	0.937	0.140	599.1	0.929	0.915	0.148	221.0
Model 2, $n = a + b\lambda_3 + c\lambda_{12} + d\rho_c$:								
Coefficients	a	b	c	d	a	b	c	d
	−0.522	−1.695	0	8.473	−0.337	0.347	1.999	13.44
Statistics	R^2	R^2_{CV}	r.m.s.	F-stat	R^2	R^2_{CV}	r.m.s.	F-stat
	0.967	0.962	0.111	494.4	0.964	0.955	0.107	292.2
Model 3, $n = a + b\nabla^2\rho c + c\rho_c + dX$:								
Coefficients	a	b	c	$d\rho^2$	a	b	c	$d\rho^2$
	−0.123	0	0	17.97	−2.109	0.799	27.05	−42.05
Statistics	R^2	R^2_{CV}	r.m.s.	F-stat	R^2	R^2_{CV}	r.m.s.	F-stat
	0.956	0.951	0.125	760.1	0.932	0.915	0.147	150.9
Model 4, $n = a + b\ G_c$:								
Coefficients	a	b			a	b		
	0.639	8.477			1.017	3.250		
Statistics	R^2	R^2_{CV}	r.m.s.	F-stat	R^2	R^2_{CV}	r.m.s.	F-stat
	0.947	0.937	0.137	584.0	0.934	0.927	0.141	491.9
Pauling, $n = \exp[(r_0 - r)/a]$:								
Coefficients	R_0	a			R_0	a		
	1.521	0.293			1.865	0.363		
Statistics	R^2	r.m.s.	F-stat		R^2	r.m.s.	F-stat	
	0.960	0.119	844.9		0.953	0.118	711.3	
Bader, $n = \exp[(r_0 - r)/a]$:								
Coefficients	b	a			b	a		
	0.243	0.151			0.163	0.103		
Statistics	R^2	r.m.s.	F-stat		R^2	r.m.s.	F-stat	
	0.955	0.127	741.3		0.863	0.202	221.4	

provide the most rigorous test possible for the various bond order models.

Pauling bond order–bond length relationships

We first consider the performance of the usual Pauling bond order expression for the various types of bond. The fit statistics are reported in Tables 2 and 3. Figures 2(a)–6(a) are plots of observed versus calculated bond order for the five types of bonds. The main finding is that the Pauling expression is unable to consistently model the data for all bond types: C—C, C—N and C—S bonds are modelled with similar, arguably acceptable accuracy (R^2 ranging from 0.93 to 0.96), but the fit is much worse for C—O and C—P bonds. The information suggested by the statistics is also reflected in the larger degree of scatter relative to the fit line in the plots for C—O and C—P bonds. The r.m.s. error on the C—O bond of ~0.12 is particularly high considering the narrower range in bond order for this type of bond (i.e. 1–2, no triple bonds possible).

Values of the a parameter, which fixes the rate of exponential decay of bond order with respect to bond length, vary from 0.29 to 0.41 for these five types of bond all involving carbon. This broadly corroborates the values found by previous work to derive these parameters from *ab initio* data.[8] However, these values should be the most reliable reported so far for these types of bond, given the extensiveness of the data set and the good reliability level of the (gradient corrected) density functional method used.

Bader bond order–bond critical point relationships

This model of bond order has really only been tested previously on C—C bonds,[9] and the data in Tables 2 and 3 show that only for this type of bond does it provide a fit of similar quality (marginally worse, in fact) than the Pauling expression based on bond length. Indeed, the Pauling expression is generally significantly better. In the case of C—P bonds the Bader expression fails completely, since the data fall into two clusters which together cannot be fitted at all by a single exponential in the bond

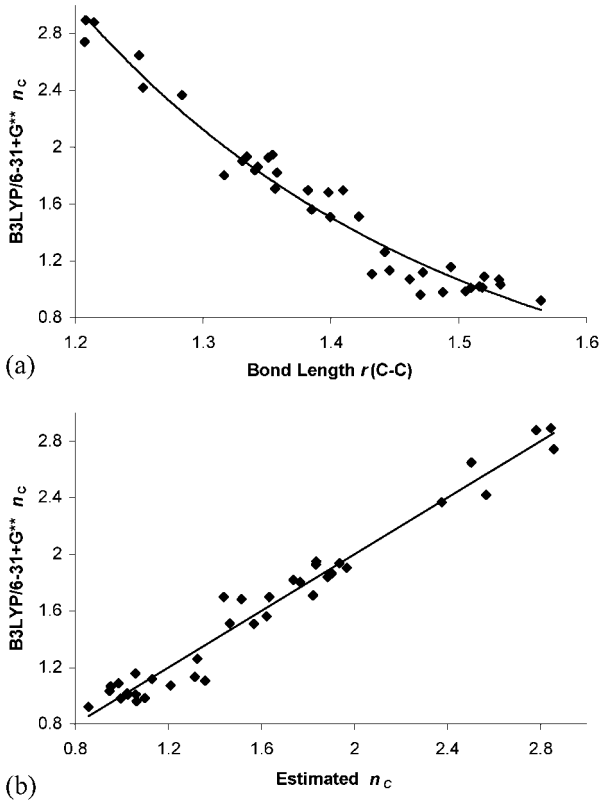

Figure 2. Plots of B3LYP/6–31 + G** Cioslowski–Mixon bond order n_C for 37 C—C bonds versus the estimated bond order for (a) the Pauling model, $n_C = \exp[(r_0 - r)/a]$, and (b) the model $n_C = a + b\,\lambda_3 + c\,\lambda_{12} + d\,\rho_c$

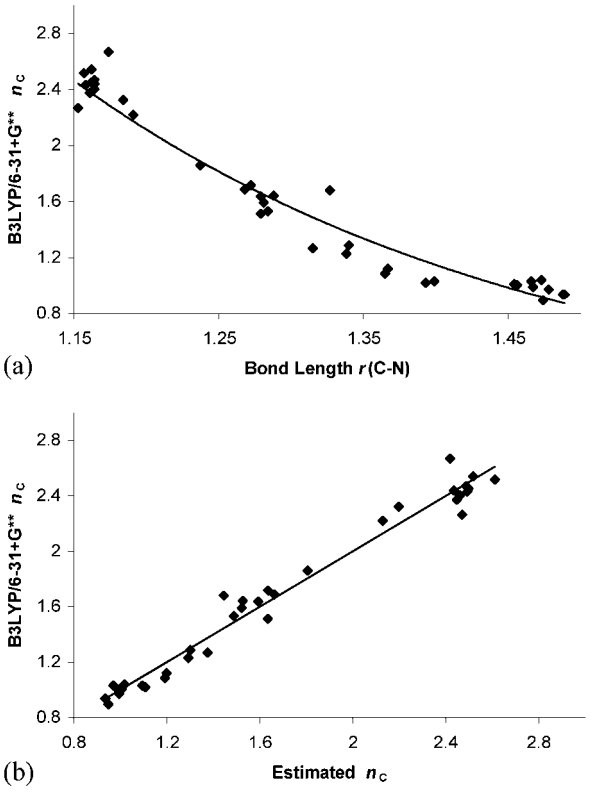

Figure 3. Plots of B3LYP/6–31 + G** Cioslowski–Mixon bond order n_C for 37 C—N bonds versus the estimated bond order for (a) the Pauling model, $n_C = \exp[(r_0 - r)/a]$, and (b) the model $n_C = a + b\,\lambda_3 + c\,\lambda_{12} + d\,\rho_c$

CP density. We can therefore discount this simple expression based on ρ_c alone as being insufficiently general.

The matrices of linear correlation coefficients linking the various descriptors are given in full in the Electronic Supplementary Information (Table 6s). The data show that ρ_c and bond length are strongly linearly correlated in four of the five bond types, the exception being the C—P bond. This explains how the two models (Pauling and Bader) can perform so differently in this particular case.

Improved bond critical point topological models

The first step was to consider whether including the isotropic measure of charge concentration/depletion $\nabla^2\rho_c$ can provide improved models of bond order based on bond CP properties. The correlation coefficients linking $\nabla^2\rho_c$ and ρ_c are low (between -0.7 and -0.8), with the exception of the non-polar C—C bonds (correlation coefficient -0.97). This verifies that these two variables are reasonably independent of one another and therefore potentially contain different information relating to bond order. Examination of the statistics for the two-parameter model $n_C = a + b\nabla^2\rho_c + c\rho_c$ (model 1 in Tables 2 and 3) shows at once that this offers a considerable improvement over the Bader model, and in three of the five cases it is also more effective than the Pauling expression. The most striking improvement compared with the Bader and Pauling models is seen for the C—O data, for which both of these expressions gave poor models. The C—P data are also modelled much better by this new equation, although the coefficient c of ρ_c in this case is zero within the estimated error, so the correlation is effectively just with $\nabla^2\rho_c$. In the case of C—C bonds, where this model gives a poorer fit than the Pauling model, a similar effect is observed except that this time it is the coefficient of $\nabla^2\rho_c$ which is zero. This can be understood in the context of the high linear correlation coefficient linking $\nabla^2\rho_c$ and ρ_c in this instance: apparently $\nabla^2\rho_c$ contains little new information compared with ρ_c for such non-polar bonds, and is not needed.

Given the moderate success of this model, it seems appropriate to consider the separation of $\nabla^2\rho_c$ into its components as discussed previously. Again, the correlation coefficients linking ρ_c, λ_{12} and λ_3 suggest that ρ_c is not correlated any more closely with the separated hessian eigenvalues than with $\nabla^2\rho_c$ itself. This generates a three-parameter model $n_C = a + b\lambda_3 + c\lambda_{12} + d\rho_c$.

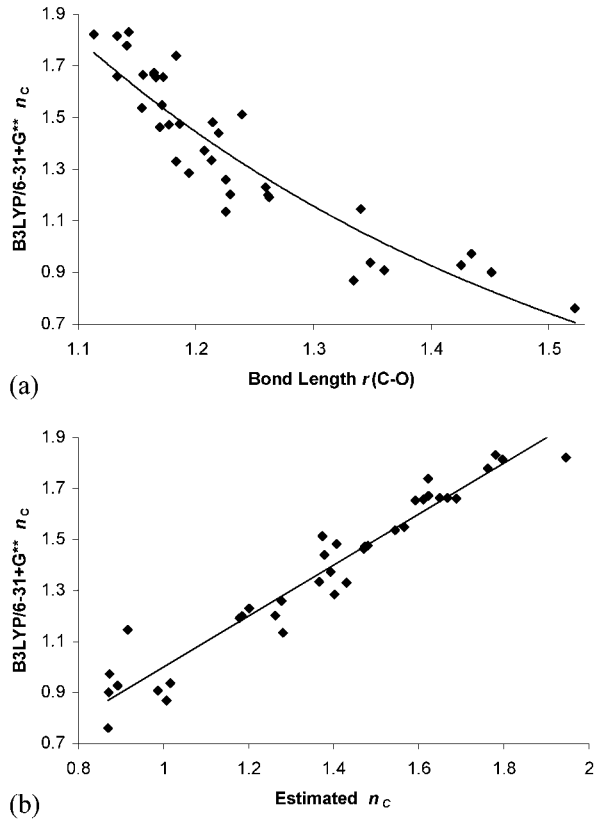

Figure 4. Plot of B3LYP/6–31 + G** Cioslowski–Mixon bond order n_C for 37 C—O bonds versus the estimated bond order for (a) the Pauling model, $n_C = \exp[(r_0 - r)/a]$, and (b) the model $n_C = a + b\,\lambda_3 + c\,\lambda_{12} + d\,\rho_c$

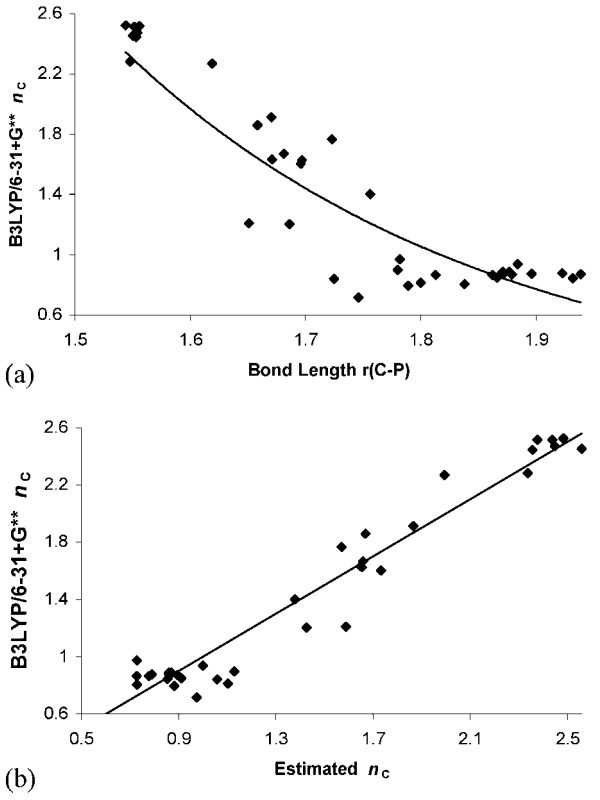

Figure 5. Plot of B3LYP/6–31 + G** Cioslowski–Mixon bond order n_C for 37 C—P bonds versus the estimated bond order for (a) the Pauling model, $n_C = \exp[(r_0 - r)/a]$, and (b) the model $n_C = a + b\,\lambda_3 + c\,\lambda_{12} + d\,\rho_c$

Again, the statistics (model 2 in Tables 2 and 3) reveal a marked improvement in the performance of this model (now better than Pauling in every case). Despite now having three parameters the *F*-stat statistic, with the exception of the model for C—P, is still fairly competitive with the values found for one- and two-parameter models. Plots of exact versus predicted bond orders are presented in Figs 2(b)–6(b), which all show the desired linear relationships.

As an alternative to separating $\nabla^2\rho_c$, we have explored models in which an additional quadratic term in either ρ_c or $\nabla^2\rho_c$ is added to two-parameter model, to produce $n_C = a + b\nabla^2\rho_c + c\rho_c + dX$, where $X = \rho_c{}^2$ or $(\nabla^2\rho_c)^2$. The statistics for this model 3 in Tables 2 and 3 show that this is marginally more effective than model 2 for bonds between carbon and the first-row atoms, but poorer than model 2 for bonds between carbon and the second-row atoms sulphur and phosphorus. We therefore suggest that model 2 is the most general and effective bond CP model of the covalent bond order.

We also briefly considered whether the Pauling models can be significantly improved by adding a term linear in $\nabla^2\rho_c$ (the generally low linear correlation coefficients linking bond length and $\nabla^2\rho_c$ suggest that this is a possibility). The answer to this question is no: the best case is for C—O bonds where the expression

$$n_C \approx \exp[(1.45 - r(\text{C—O})_{opt})/0.84] + 0.26\nabla^2\rho_c$$
$$(R^2 = 0.91, \text{r.m.s. error} = 0.10)$$

does provide a considerable improvement over the basic Pauling expression. However, since even this two-descriptor model is poorer than the $\{\rho_c, \nabla^2\rho_c\}$ two-descriptor model, we can discount this particular combination of parameters as not being very generally useful.

Models based on the bond critical point kinetic energy density

In order to gauge whether G_c is a suitable descriptor of covalent bond order, we consider the simple linear correlation $n_C = a + b\,G_c$ (model 4 in Tables 2 and 3). The results show that in some types of bond it is a remarkably good descriptor of bond order. In the case of the C—N bond (Fig. 7), this simple one-parameter linear model in G_c is competitive with our best and most general, three-parameter model (model 2). It also shows a good linear correlation with C—O bond order, far better than the Pauling or Bader models in this case. The overall

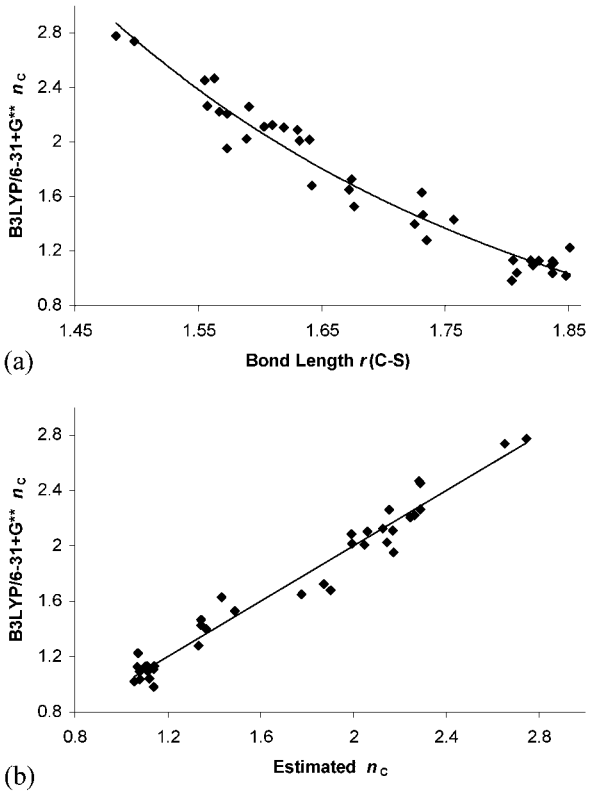

Figure 6. Plot of B3LYP/6–31 + G** Cioslowski–Mixon bond order n_C for 37 C—S bonds versus the estimated bond order for (a) the Pauling model, $n_C = \exp[(r_0 - r)/a]$, and (b) the model $n_C = a + b\,\lambda_3 + c\,\lambda_{12} + d\,\rho_c$

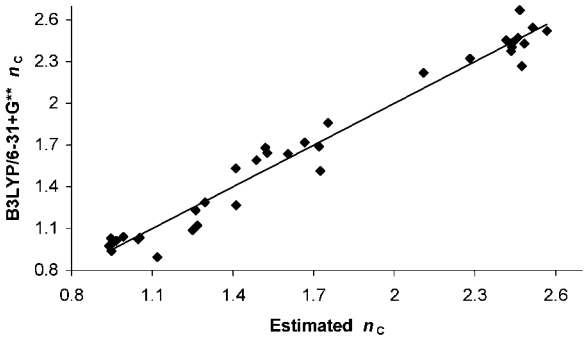

Figure 7. Plot of B3LYP/6–31 + G** Cioslowski–Mixon bond order n_C for 37 C—N bonds versus the estimated bond order for the model $n_C = a + b\,G_c$

performance suggests that it is marginally better than bond length as a single-parameter description of bond order.

Whereas the linear correlation between G_c and bond length is high (except for non-polar C—C), the lesser correlations between G_c and ρ_c suggest that these parameters might be useful in combination. Various linear and non-linear models with combinations of G_c and bond length or ρ_c were tried, but except for a few isolated cases in particular types of bond, no significant improvement was obtained.

Finally, we note that correlations of bond order with the alternative form of kinetic energy density[30] $K_c = G_c - 1/4\nabla^2\rho_c$ were also attempted. These were found to be much poorer than those for G_c.

CONCLUSIONS

Analyses of covalent bond orders calculated using the Cioslowski–Mixon equation for a range of bond types broadly supports the historical use of the Pauling relationship as a simple, one-parameter description of bond order. In polar bonds it would appear that a simple linear model in the bond critical point kinetic energy density correlates even more closely with the actual (covalent) bond order than the Pauling bond length-based expression. However, it must be emphasized that neither the bond length nor any single charge density topological parameter acts as an effective, unified descriptor of bond order for all of the five bond types considered here. The simple exponential relationship between bond order and the bond critical point electron density suggested by Bader seems to be restricted to the case of bonds between carbon atoms.

Studying the interdependences of bond length and various bond charge density topological parameters reveals some interesting, hitherto unrecognized, results. Undoubtedly the bonds between carbon and phosphorus show the most atypical behaviour—bond length is not closely correlated with bond critical point electron density, and perhaps because of this the Pauling bond order–bond length expression is particularly poor in this case. The key finding is that the bond perpendicular and bond parallel principal curvatures of the electron density at the bond critical point are able to act as reasonably independent descriptors of bond order. This in turn enables simple charge density topological models to be constructed which correlate with covalent bond order far more closely than the Pauling relationship.

The multiple linear description $n_C = a + b\ \rho_c + c\ \lambda_3 + d\ (\lambda_1 + \lambda_2)$ is recommended since it works well for all five multiple bonds studied and not just for neutral, closed-shell species (the data analysed contained plenty of examples of anions and cations, plus neutral and charged radicals). It has an elegant and simple physical interpretation: ρ_c and λ_3 measure σ character, whilst the curvatures perpendicular to the bond $(\lambda_1 + \lambda_2)$ measure the degree of π character. It is the case that not every term in the expression is required to describe bond order for every bond type: non-polar C—C bonds, for example, do not require a $(\lambda_1 + \lambda_2)$ term. However, it appears that this expression is sufficiently flexible to provide a good model for both polar and non-polar bonds.

It is therefore possible to estimate rigorously the covalent part of the bond order from a topological

analysis of the charge distribution for large molecules where the atomic overlap matrix methods become impractical to apply. It is also possible to estimate bond orders from a topological analysis of experimentally derived charge distributions (from elastic x-ray scattering) using these techniques.

Acknowledgements

The authors are grateful to Dr J. A. Platts for helpful comments and a critical reading of the manuscript.

REFERENCES

1. Pauling L. *The Nature of the Chemical Bond*. Cornell University Press: Ithaca, NY, 1960; 144.
2. Lendvay G. *J. Phys. Chem.* 1994; **98**: 6098–6104.
3. Shustorovich E. *Surf. Sci.* 1985; **163**: L645–654.
4. Blowers P, Masel RI. *J. Phys. Chem. A* 1998; **102**: 9957–9964.
5. Bridgeman AJ, Cavigliasso G, Ireland LR, Rothery J. *J. Chem. Soc. Dalton Trans.* 2001; 2095–2108.
6. Krygowski TM, Ciesielski A, Bird CW. *J. Chem. Inf. Comput. Sci.* 1995; **35**: 203–210.
7. Nalewajski RF. *J. Phys. Chem. A* 2000; **104**: 11940–11951.
8. Lendvay G. *J. Mol. Struct. Theochem.* 2000; **501**: 389–393.
9. Bader RFW. *Atoms In Molecules –A Quantum Theory*. Oxford University Press: Oxford, 1990; 252.
10. Silvi B, Savin A. *Nature (London)* 1994; **371**: 683–686.
11. Wiberg KA. *Tetrahedron* 1968; **24**: 1083–1089.
12. Mayer I. *Chem. Phys. Lett.* 1983; **97**: 270–274.
13. Cioslowski J, Mixon ST. *J. Am. Chem. Soc.* 1991; **113**: 4142–4145.
14. Biegler-König FW, Bader RFW, Tang TH. *J. Comput. Chem.* 1982; **3**: 317–328.
15. Angyan J, Loos M, Mayer I. *J. Phys. Chem.* 1994; **98**: 5244–5248.
16. Boyd RJ, Edgecombe KE. *J. Am. Chem. Soc.* 1988; **110**: 4182–4186.
17. Grimme S. *J. Am. Chem. Soc.* 1996; **118**: 1529–1534.
18. Exner K, Schleyer PvR. *J. Phys. Chem. A* 2001; **105**: 3407–3416.
19. Howard ST, Krygowski TM. *Can. J. Chem.* 1997; **75**: 1174–1181.
20. Boyd RJ, Choi SC. *Chem. Phys. Lett.* 1985; **120**: 80–85.
21. Platts JA. *Phys. Chem. Chem. Phys.* 2000; 973–980.
22. Popelier PLA. *J. Phys. Chem. A* 1999; **103**: 2883–2890.
23. Coppens P. *X-Ray Charge Densities And Chemical Bonding*. Oxford University Press: Oxford, 1997.
24. Becke AD. *Phys. Rev. A* 1988; **38**: 3098–3100.
25. Lee C, Yang W, Parr RG. *Phys. Rev. B* 1988; **41**: 785–789.
26. (a) Mclean AD, Chandler GS. *J. Chem. Phys.* 1980; **72**: 5639–5645; (b) Krishnan R, Binkley JS, Seeger R, Pople JA. *J. Chem. Phys.* 1980; **72**: 650–656.
27. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA Jr., Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Baboul AG, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Andres JL, Gonzalez C, Head-Gordon M, Replogle ES, Pople JA. Gaussian 98. Gaussian: Pittsburgh, PA, 1998.
28. JMP Version 4.0.2. SAS Institute: Cary, NC, USA, 2000.
29. Wolfram S. Mathematica Book, 4th edition, Wolfram media/ Cambridge University Press, 1999.
30. Bader RFW, Preston HJT. *Int. J. Quantum Chem.* 1969; **3**: 327–339.