Generalized Gradient Approximation Made Simple

John P. Perdew, Kieron Burke,* Matthias Ernzerhof
Department of Physics and Quantum Theory Group, Tulane University, New Orleans, Louisiana 70118
(Received 21 May 1996)

Generalized gradient approximations (GGA's) for the exchange-correlation energy improve upon the local spin density (LSD) description of atoms, molecules, and solids. We present a simple derivation of a simple GGA, in which all parameters (other than those in LSD) are fundamental constants. Only general features of the detailed construction underlying the Perdew-Wang 1991 (PW91) GGA are invoked. Improvements over PW91 include an accurate description of the linear response of the uniform electron gas, correct behavior under uniform scaling, and a smoother potential. [S0031-9007(96)01479-2]

PACS numbers: 71.15.Mb, 71.45.Gm

Kohn-Sham density functional theory [1,2] is widely used for self-consistent-field electronic structure calculations of the ground-state properties of atoms, molecules, and solids. In this theory, only the exchange-correlation energy $E_{\mathrm{XC}} = E_{\mathrm{X}} + E_{\mathrm{C}}$ as a functional of the electron spin densities $n_\uparrow(\mathbf{r})$ and $n_\downarrow(\mathbf{r})$ must be approximated. The most popular functionals have a form appropriate for slowly varying densities: the local spin density (LSD) approximation

$$E_{\mathrm{XC}}^{\mathrm{LSD}}[n_\uparrow, n_\downarrow] = \int d^3r\, n\epsilon_{\mathrm{XC}}^{\mathrm{unif}}(n_\uparrow, n_\downarrow), \qquad (1)$$

where $n = n_\uparrow + n_\downarrow$, and the generalized gradient approximation (GGA) [3,4]

$$E_{\mathrm{XC}}^{\mathrm{GGA}}[n_\uparrow, n_\downarrow] = \int d^3r\, f(n_\uparrow, n_\downarrow, \nabla n_\uparrow, \nabla n_\downarrow). \qquad (2)$$

In comparison with LSD, GGA's tend to improve total energies [4], atomization energies [4–6], energy barriers and structural energy differences [7–9]. GGA's expand and soften bonds [6], an effect that sometimes corrects [10] and sometimes overcorrects [11] the LSD prediction. Typically, GGA's favor density inhomogeneity more than LSD does.

To facilitate practical calculations, $\epsilon_{\mathrm{XC}}^{\mathrm{unif}}$ and f must be parametrized analytic functions. The exchange-correlation energy per particle of a uniform electron gas, $\epsilon_{\mathrm{XC}}^{\mathrm{unif}}(n_\uparrow, n_\downarrow)$, is well established [12], but the best choice for $f(n_\uparrow, n_\downarrow, \nabla n_\uparrow, \nabla n_\downarrow)$ is still a matter of debate. Judging the derivations and formal properties of various GGA's can guide a rational choice among them. Semiempirical GGA's can be remarkably successful for small molecules, but fail for delocalized electrons in the uniform gas [when $f(n_\uparrow, n_\downarrow, 0, 0) \neq n\epsilon_{\mathrm{XC}}^{\mathrm{unif}}(n_\uparrow, n_\downarrow)$] and thus in simple metals. A first-principles numerical GGA can be constructed [13] by starting from the second-order density-gradient expansion for the exchange-correlation hole surrounding the electron in a system of slowly varying density, then cutting off its spurious long-range parts to satisfy sum rules on the exact hole. The Perdew-Wang 1991 (PW91) [14] functional is an analytic fit to this numerical GGA, designed to satisfy several further exact conditions [13]. PW91 incorporates some inhomogeneity effects while retaining many of the best features of LSD, but has its own problems: (1) The derivation is long, and depends on a mass of detail. (2) The analytic function f, fitted to the numerical results of the real-space cutoff, is complicated and nontransparent. (3) f is overparametrized. (4) The parameters are not seamlessly joined [15], leading to spurious wiggles in the exchange-correlation potential $\delta E_{\mathrm{XC}}/\delta n_\sigma(\mathbf{r})$ for small [16] and large [16,17] dimensionless density gradients, which can bedevil the construction of GGA-based electron-ion pseudopotentials [18–20]. (5) Although the numerical GGA correlation energy functional behaves properly [13] under Levy's uniform scaling to the high-density limit [21], its analytic parametrization (PW91) does not [22]. (6) Because PW91 reduces to the second-order gradient expansion for density variations that are either slowly varying *or* small, it descibes the linear response of the density of a uniform electron gas *less* satisfactorily than does LSD [20,23].

This last problem illustrates a fact which is often overlooked: The semilocal form of Eq. (2) is too restrictive to reproduce all the known behaviors of the exact functional [13]. In contrast to the construction of the PW91 functional, which was designed to satisfy as many exact conditions as possible, the GGA presented here satisfies only those which are energetically significant. For example, in the pseudopotential theory of simple metals, the linear-response limit is physically important. On the other hand, recovery of the exact second-order gradient expansion in the slowly varying limit makes little difference to the energies of real systems. We solve the 6 problems above with a simple new derivation of a simple new GGA functional in which *all* parameters [other than those in $\epsilon_{\mathrm{XC}}^{\mathrm{unif}}(n_\uparrow, n_\downarrow)$] are fundamental constants. Although the derivation depends only on the most general features of the real-space construction [13] behind PW91, the resulting functional is close to numerical GGA.

We begin with the GGA for correlation in the form

$$E_{\mathrm{C}}^{\mathrm{GGA}}[n_\uparrow, n_\downarrow] = \int d^3r\, n[\epsilon_{\mathrm{C}}^{\mathrm{unif}}(r_s, \zeta) + H(r_s, \zeta, t)], \qquad (3)$$

where r_s is the local Seitz radius ($n = 3/4\pi r_s^3 = k_F^3/3\pi^2$), $\zeta = (n_\uparrow - n_\downarrow)/n$ is the relative spin polarization, and $t = |\nabla n|/2\phi k_s n$ is a dimensionless density gradient [13,14]. Here $\phi(\zeta) = [(1+\zeta)^{2/3} + (1-\zeta)^{2/3}]/2$ is a spin-scaling factor [24], and $k_s = \sqrt{4k_F/\pi a_0}$ is the Thomas-Fermi screening wave number ($a_0 = \hbar^2/me^2$). $\nabla\zeta$ corrections to Eq. (3), which are small for most purposes, will be derived in later work. We construct the gradient contribution H from three conditions:

(a) In the slowly varying limit ($t \to 0$), H is given by its second-order gradient expansion [24]

$$H \to (e^2/a_0)\beta\phi^3 t^2, \qquad (4)$$

where $\beta \simeq 0.066\,725$. This is the high-density ($r_s \to 0$) limit [25] of the weakly r_s-dependent gradient coefficient [26] for the correlation energy [with a Yukawa interaction $(e^2/u)\exp(-\kappa u)$ in the limit $\kappa \to 0$], and also the coefficient which emerges naturally from the numerical GGA [13] discussed earlier.

(b) In the rapidly varying limit $t \to \infty$,

$$H \to -\epsilon_C^{\rm unif}, \qquad (5)$$

making correlation vanish. As $t \to \infty$ in the numerical GGA, the sum rule $\int d^3u\, n_C(\mathbf{r}, \mathbf{r}+\mathbf{u}) = 0$ on the correlation hole density n_C is only satisfied by $n_C = 0$. For example, in the tail of the electron density of a finite system, the exchange energy density and potential dominate their correlation counterparts in reality, but not in LSD.

(c) Under uniform scaling to the high-density limit [$n(\mathbf{r}) \to \lambda^3 n(\lambda\mathbf{r})$ and $\lambda \to \infty$, whence $r_s \to 0$ as λ^{-1} and $t \to \infty$ as $\lambda^{1/2}$], the correlation energy must scale to a constant [21]. Thus [27] H must cancel the logarithmic singularity of $\epsilon_C^{\rm unif}$ [28] in this limit: $\epsilon_C^{\rm unif}(r_s, \zeta) \to (e^2/a_o)\phi^3[\gamma\ln(r_s/a_0) - \omega]$, where γ and ω are weak functions [12] of ζ which we shall replace by their $\zeta = 0$ values, $\gamma = (1 - \ln 2)/\pi^2 \simeq 0.031\,091$ and $\omega \simeq 0.046\,644$, so

$$H \to (e^2/a_0)\gamma\phi^3 \ln t^2. \qquad (6)$$

Conditions (a), (b), and (c) are satisfied by the simple ansatz

$$H = (e^2/a_0)\gamma\phi^3 \times \ln\left\{1 + \frac{\beta}{\gamma} t^2 \left[\frac{1 + At^2}{1 + At^2 + A^2t^4}\right]\right\}, \qquad (7)$$

where

$$A = \frac{\beta}{\gamma}\left[\exp\{-\epsilon_C^{\rm unif}/(\gamma\phi^3 e^2/a_0)\} - 1\right]^{-1}. \qquad (8)$$

The function H starts out from $t = 0$ like Eq. (4), and grows monotonically to the limit of Eq. (5) as $t \to \infty$; thus $E_C^{\rm GGA} \le 0$. (H appears as one of two terms in the PW91 correlation energy, but with $\gamma = 0.025$.) Under uniform scaling to the high density limit, $E_C^{\rm GGA}$ tends to

$$-\frac{e^2}{a_0}\int d^3r\, n\gamma\phi^3 \times \ln\left[1 + \frac{1}{\chi s^2/\phi^2 + (\chi s^2/\phi^2)^2}\right], \qquad (9)$$

where $s = |\nabla n|/2k_F n = (r_s/a_0)^{1/2}\phi t/c$ is another dimensionless density gradient, $c = (3\pi^2/16)^{1/3} \simeq 1.2277$, and $\chi = (\beta/\gamma)c^2\exp(-\omega/\gamma) \simeq 0.721\,61$. The correlation energy for a two-electron ion of nuclear charge $Z \to \infty$ is $-\infty$ by LSD, $+\infty$ by PW91, -0.0482 Hartree by Eq. (9), and -0.0467 exactly [29]. For a *finite* system, s cannot vanish identically, except on sets of measure zero, so Eq. (9) is finite; for an *infinite* jellium, s vanishes everywhere, and Eq. (9) reduces to $-\infty$ as GGA reduces to LSD.

The GGA for the exchange energy will be constructed from four further conditions:

(d) Under the uniform density scaling described along with condition (c) above, E_X must scale [30] like λ. Thus, for $\zeta = 0$ everywhere, we must have

$$E_X^{\rm GGA} = \int d^3r\, n\epsilon_X^{\rm unif}(n)F_X(s), \qquad (10)$$

where $\epsilon_X^{\rm unif} = -3e^2k_F/4\pi$. To recover the correct uniform gas limit, $F_X(0) = 1$.

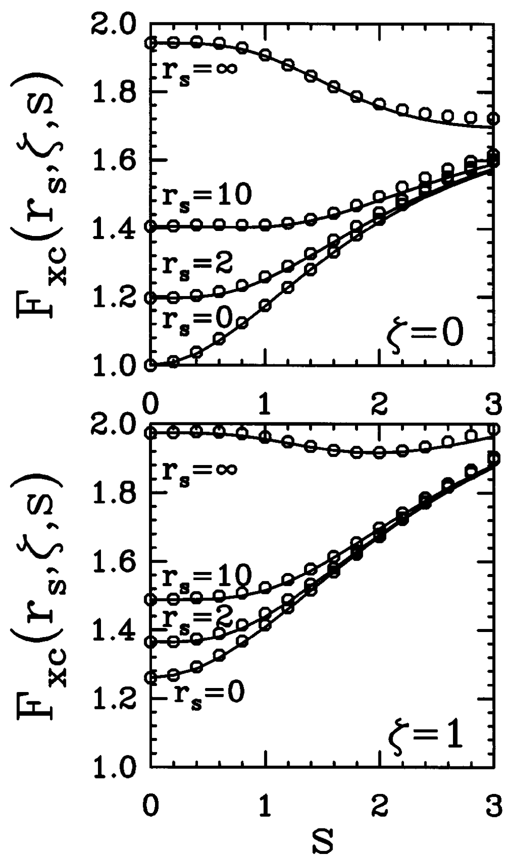

FIG. 1. Enhancement factors of Eq. (15) showing GGA nonlocality. Solid curves denote the present GGA, while open circles denote the PW91 of Refs. [4,13,14].

(e) The exact exchange energy obeys the spin-scaling relationship [31]

$$E_X[n_\uparrow, n_\downarrow] = (E_X[2n_\uparrow] + E_X[2n_\downarrow])/2\,. \qquad (11)$$

(f) For the linear response of the spin-unpolarized uniform electron gas, i.e., for small density variations around the uniform density, LSD is an excellent approximation [23] to the exchange-correlation energy, while the gradient expansion is not [20]. To recover the LSD linear response, we must have (as $s \to 0$)

$$F_X(s) \to 1 + \mu s^2, \qquad (12)$$

where $\mu = \beta(\pi^2/3) \simeq 0.219\,51$, the effective gradient coefficient for exchange, cancels that for correlation.

(g) The Lieb-Oxford bound [14]

$$E_X[n_\uparrow, n_\downarrow] \geq E_{XC}[n_\uparrow, n_\downarrow] \geq -1.679e^2 \int d^3r\, n^{4/3} \qquad (13)$$

will be satisfied if the spin-polarized enhancement factor $F_X(\zeta = 1, s) = 2^{1/3}F_X(s/2^{1/3})$ grows gradually with s to a maximum value less than or equal to 2.273, i.e., $F_X(s) \leq 1.804$. A simple $F_X(s)$ satisfying Eqs. (12) and (13) is

$$F_X(s) = 1 + \kappa - \kappa/(1 + \mu s^2/\kappa), \qquad (14)$$

where $\kappa = 0.804$. Becke [32] proposed this form, but with empirical coefficients ($\kappa = 0.967, \mu = 0.235$).

To portray the nonlocality of the new GGA, we define the enhancement factor [4] F_{XC} over local exchange

$$E_{XC}^{GGA}[n_\uparrow, n_\downarrow] = \int d^3r\, n\epsilon_X^{unif}(n)F_{XC}(r_s, \zeta, s). \qquad (15)$$

Equation (15) represents any GGA exactly when ζ is independent of **r**, and is always approximately valid. LSD corresponds to the further approximation $F_{XC}(r_s, \zeta, s) \to F_{XC}(r_s, \zeta, 0)$. Figure 1 displays the nonlocality or s dependence of F_{XC} for $\zeta = 0$ and 1 for our simplified GGA and for PW91, demonstrating their numerical similarity. The range of interest for real systems [9] is $0 \lesssim s \lesssim 3$ and $0 \lesssim r_s/a_0 \lesssim 10$. Exchange dominates the high-density limit, with $F_{XC}(r_s, \zeta, s) \to F_X(\zeta, s)$ as $r_s \to 0$, where $F_X(\zeta, s)$ is found from Eqs. (10) and (11) by dropping $\nabla\zeta$ contributions. As the reduced gradient s increases at fixed r_s, exchange turns on more strongly, while correlation turns off. The exchange nonlocality is dominant for valence-electron densities ($1 \lesssim r_s/a_0 \lesssim 10$), so that for most physical r_s GGA favors density inhomogeneity

TABLE I. Atomization energies of molecules, in kcal/mol (1 eV = 23.06 kcal/mol). E_{XC} has been evaluated on self-consistent densities at experimental geometries [33]. Nonspherical densities and Kohn-Sham potentials have been used for open-shell atoms [34]. The calculations are performed with a modified version of the CADPAC program [35]. The experimental values for ΔE (with zero point vibration removed) are taken from Ref. [36]. PBE is the simplified GGA proposed here. UHF is unrestricted Hartree-Fock, for comparison.

System	ΔE^{UHF}	ΔE^{LSD}	ΔE^{PW91}	ΔE^{PBE}	ΔE^{expt}
H_2	84	113	105	105	109
LiH	33	60	53	52	58
CH_4	328	462	421	420	419
NH_3	201	337	303	302	297
OH	68	124	110	110	107
H_2O	155	267	235	234	232
HF	97	162	143	142	141
Li_2	3	23	20	19	24
LiF	89	153	137	136	139
Be_2	−7	13	10	10	3
C_2H_2	294	460	415	415	405
C_2H_4	428	633	573	571	563
HCN	199	361	326	326	312
CO	174	299	269	269	259
N_2	115	267	242	243	229
NO	53	199	171	172	153
O_2	33	175	143	144	121
F_2	−37	78	54	53	39
P_2	36	142	120	120	117
Cl_2	17	81	64	63	58
Mean abs. error	71.2	31.4	8.0	7.9	···

more than LSD does. Only in the low-density limit is correlation so strong in comparison with exchange that the correlation nonlocality dominates. Note also the approximate cancellation of nonlocality in the low-density fully spin-polarized ($r_s \to \infty, \zeta = 1$) limit, where the exchange-correlation hole becomes as deep, and therefore (by the sum rule [13]) as local, as it can be on the scale set by the local density itself.

The GGA proposed here retains correct features of LSD, and combines them with the most energetically important features of gradient-corrected nonlocality. The correct but less important features of PW91 which have been sacrificed are (1) correct second-order gradient coefficients for E_X and E_C in the slowly varying limit, and (2) correct nonuniform scaling of E_X in limits where the reduced gradient s tends to ∞.

Calculations of atomization energies for small molecules (Table I) also show that our simple new E_{XC}^{GGA} yields essentially the same results as the more Byzantine PW91. Except for its additional satisfaction of conditions (c) and (f) and its smoother potential, the new GGA is close to PW91. However, its simpler form and derivation make it easier to understand, apply, and perhaps improve.

J. P. P. acknowledges discussion of the linear-response limit with Raffaele Resta. We thank Don Hamann, Cyrus Umrigar, and Claudia Filippi for suggestions and tests. This work was supported by the National Science Foundation under Grant No. DMR95-21353, and in part by the Deutsche Forschungsgemeinschaft.

*Permanent address: Department of Chemistry, Rutgers University, Camden, NJ 08102.

[1] W. Kohn and L. J. Sham, Phys. Rev. **140**, A 1133 (1965).

[2] R. M. Dreizler and E. K. U. Gross, *Density Functional Theory* (Springer-Verlag, Berlin, 1990); R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules* (Oxford, New York, 1989).

[3] D. C. Langreth and M. J. Mehl, Phys. Rev. B **28**, 1809 (1983); A. D. Becke, Phys. Rev. A **38**, 3098 (1988).

[4] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992); **48**, 4978(E) (1993).

[5] A. D. Becke, J. Chem. Phys. **96**, 2155 (1992).

[6] E. I. Proynov, E. Ruiz, A. Vela, and D. R. Salahub, Int. J. Quantum Chem. **S29**, 61 (1995); A. C. Scheiner, J. Baker, and J. W. Andzelm (unpublished).

[7] B. Hammer, K. W. Jacobsen, and J. K. Nørskov, Phys. Rev. Lett. **70**, 3971 (1993); B. Hammer and M. Scheffler, Phys. Rev. Lett. **74**, 3487 (1995).

[8] D. R. Hamann, Phys. Rev. Lett. **76**, 660 (1996); P. H. T. Philipsen, G. te Velde, and E. J. Baerends, Chem. Phys. Lett. **226**, 583 (1994).

[9] A. Zupan, J. P. Perdew, K. Burke, and M. Causá, Int. J. Quantum Chem. (to be published).

[10] V. Ozolins and M. Körling, Phys. Rev. B **48**, 18 304 (1993).

[11] C. Filippi, D. J. Singh, and C. Umrigar, Phys. Rev. B **50**, 14 947 (1994).

[12] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13 244 (1992), and references therein.

[13] J. P. Perdew, K. Burke, and Y. Wang, Phys. Rev. B (to appear).

[14] J. P. Perdew, in *Electronic Structure of Solids '91*, edited by P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991), p. 11.

[15] E. Engel and S. H. Vosko, Phys. Rev. B **47**, 13 164 (1993).

[16] C. Filippi, C. J. Umrigar, and M. Taut, J. Chem. Phys. **100**, 1290 (1994).

[17] R. Neumann, R. H. Nobes, and N. C. Handy, Mol. Phys. **87**, 1 (1996).

[18] G. Ortiz and P. Ballone, Phys. Rev. B **43**, 6376 (1991).

[19] R. N. Barnett and U. Landman, Phys. Rev. Lett. **70**, 1775 (1993).

[20] G. Ortiz, Phys. Rev. B **45**, 11 328 (1992).

[21] M. Levy, Int. J. Quantum Chem. **S23**, 617 (1989).

[22] C. J. Umrigar and X. Gonze, in *High Performance Computing and its Application to the Physical Sciences, Proceedings of the Mardi Gras 1993 Conference*, edited by D. A. Browne *et al.* (World Scientific, Singapore, 1993).

[23] C. Bowen, G. Sugiyama, and B. J. Alder, Phys. Rev. B **50**, 14 838 (1994); S. Moroni, D. M. Ceperley, and G. Senatore, Phys. Rev. Lett. **75**, 689 (1995).

[24] Y. Wang and J. P. Perdew, Phys. Rev. B **43**, 8911 (1991).

[25] S.-K. Ma and K. A. Brueckner, Phys. Rev. **165**, 18 (1968).

[26] D. J. W. Geldart and M. Rasolt, Phys. Rev. B **13**, 1477 (1976); D. C. Langreth and J. P. Perdew, Phys. Rev. B **21**, 5469 (1980).

[27] M. Levy and J. P. Perdew, Phys. Rev. B **48**, 11 638 (1993).

[28] M. Gell-Mann and K. A. Brueckner, Phys. Rev. **106**, 364 (1957).

[29] S. Ivanov *et al.* (unpublished).

[30] M. Levy and J. P. Perdew, Phys. Rev. A **32**, 2010 (1985).

[31] G. L. Oliver and J. P. Perdew, Phys. Rev. A **20**, 397 (1979).

[32] A. D. Becke, J. Chem. Phys. **84**, 4524 (1986).

[33] D. J. DeFrees, B. A. Levi, S. K. Pollack, W. J. Hehre, J. S. Binkley, and J. A. Pople, J. Am. Chem. Soc. **101**, 4085 (1979). Geometries of NO, Cl_2 and P_2: K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure IV: Constants of Diatomic Molecules* (Van Nostrand Reinhold, New York, 1979). Geometry of Be_2: Ref. [36].

[34] F. W. Kutzler and G. S. Painter, Phys. Rev. Lett. **59**, 1285 (1987).

[35] CADPAC6: The Cambridge Analytical Derivatives Package Issue 6.0 Cambridge (1995). R. D. Amos, I. L. Alberts, J. S. Andrews, S. M. Colwell, N. C. Handy, D. Jayatilaka, P. J. Knowles, R. Kobayashi, G. J. Laming, A. M. Lee, P. E. Maslen, C. W. Murray, P. Palmieri, J. E. Rice, J. Sanz, E. D. Simandiras, A. J. Stone, M.-D. Su, and D. J. Tozer.

[36] J. A. Pople, M. Head-Gordon, D. J. Fox, K. Raghavachari, and L. A. Curtiss, J. Chem. Phys. **90**, 5622 (1989); L. A. Curtiss, C. Jones, G. W. Trucks, K. Raghavachari, and J. A. Pople, J. Chem. Phys. **93**, 2537 (1990). Be_2: V. E. Bondybey and J. H. English, J. Chem. Phys. **80**, 568 (1984).

ERRATA

Generalized Gradient Approximation Made Simple [Phys. Rev. Lett. 77, 3865 (1996)]

John P. Perdew, Kieron Burke, and Matthias Ernzerhof

[S0031-9007(97)02302-8]

For the molecules Be_2, F_2, and P_2 of Table I, the unrestricted Hartree-Fock solution breaks the singlet spin symmetry, even though the density-functional solutions do not. For these broken-symmetry solutions, the UHF atomization energies become $+7$, -20, and $+41$ kcal/mol, respectively, and the mean absolute error of all the UHF atomization energies becomes 69.8 kcal/mol.

The PBE correlation energy of the two-electron ions of nuclear charge $Z \to \infty$ should be corrected to -0.0479 hartree, consistent with the PBE value $\omega = 0.046644$ stated in the Letter. The quoted value -0.0482 hartree was obtained from the more refined $\omega = 0.046920$ of G. G. Hoffman, Phys. Rev. B **45**, 8730 (1992).

Reference [6] should have been "A. C. Scheiner, J. Baker, and J. W. Andzelm, J. Comput. Chem. (to be published)".

 0031-9007/97/78(7)/1396(1)$10.00