Origin-Independent Sum Over States Simulations of Magnetic and Electronic Circular Dichroism Spectra via the Localized Orbital/Local Origin Method

Petr Štěpánek[a,b] and Petr Bouř*[a]

Although electronic and magnetic circular dichroism (ECD, MCD) spectra reveal valuable details about molecular geometry and electronic structure, quantum-chemical simulations significantly facilitate their interpretation. However, the simulated results may depend on the choice of coordinate origin. Previously (Štěpánek and Bouř, J. Comput. Chem. 2013, 34, 1531), the sum-over-states (SOS) methodology was found useful for efficient MCD computations. Approximate wave functions were "resolved" using time-dependent density functional theory, and the origin-dependence was avoided by placing the origin to the center of mass of the investigated molecule. In this study, a more elegant way is proposed, based on the localized orbital/local origin (LORG) formalism, and a similar approach is also applied to generate ECD intensities. The LORG-like approach yields fully origin-independent ECD and MCD spectra. The results thus indicate that the computationally relatively cheap SOS simulations open a new way of modeling molecular properties, including those involving the origin-dependent magnetic dipole moment operator. © 2015 Wiley Periodicals, Inc.

DOI: 10.1002/jcc.23845

Introduction

Optical spectroscopies are indispensable tools to study molecular structure and interactions. In particular, circular dichroism capable of detecting differences in absorption of left- and right-circularly polarized light, proved to be much more sensitive to the structure than simple absorption, for example. Electronic (natural) circular dichroism (ECD) thus revealed structural data on a plethora of systems including peptides, proteins, and nucleic acids.[1–4]

Magnetic circular dichroism (MCD) is a technique not as common as ECD; yet its applicability has been recently enhanced by the possibility to reliably simulate spectra from the first principles using publicly available software only.[5–10] Typically, molecules with conjugated π-electrons, such as porphyrins, are studied by MCD, as these provide a strong signal, whereas their ECD would be zero or small due to molecular symmetry.[11–17]

Reliable simulations of ECD and MCD intensities have to take into account the origin dependence of the results if using an approximate wave function, obtained, for example, with an incomplete basis set of atomic orbitals. A standard way of avoiding the origin dependence is to utilize the London atomic orbitals (often referred to as gauge-independent atomic orbitals, GIAO), which is applicable to optical rotation,[18,19] ECD[20,21] and MCD.[7,22,23] The *B*-term in MCD can then be obtained, for example, as the imaginary part of the Verdet constant using damped time-dependent density functional theory (TDDFT).[24] Real-time TDDFT has also been proposed for MCD simulations and found convenient, for example, for evaluating and verifying energy-weighted sum rules.[25]

However, in some circumstances, the application of GIAOs is not convenient, for example, due to a difficult implementation.[18,25,26] A special treatment is required to address the origin problem of magnetic properties for a plane wave basis set. For example, the coordinate-space grid may introduce a dependence on the origin.[27,28] Recently, we found that the sum over states (SOS) approach provides a very efficient and accurate way of generating MCD spectra when the electronic excited states obtained from TDDFT computations are used.[29–31] In test computations, we associated the Kohn–Sham determinant[32] (very close to the Hartree–Fock one[33]) with molecular ground state, and spin-adapted TDDFT single excitations with excited electronic states. This provided sufficient accuracy and enabled a direct comparison with standard response theory. The SOS way is, however, formulated for a general wave function. Within TDDFT, including the true excited state energies into the SOS formula is critical as one-electron orbitals and energies provide only very approximate results.[34–37]

---

[a] *P. Štěpánek, P. Bouř*
Group of Molecular Spectroscopy, Institute of Organic Chemistry and Biochemistry, Academy of Sciences, Flemingovo nám. 2, 16610 Prague, Czech Republic
E-mail: bour@uochb.cas.cz

[b] *P. Štěpánek*
Institute of Physics, Faculty of Mathematics and Physics, Charles University, Ke Karlovu 5, 121 16 Prague, Czech Republic

Contract grant sponsor: Czech Science Foundation; Contract grant number: 13-03978S and P208/11/0105; Contract grant sponsor: Academy of Sciences; Contract grant number: M200551205; Contract grant sponsor: Grant Agency of Charles University; Contract grant number: 711012

SOS-like expressions for both electric and magnetic static properties within TDDFT have been explored previously, and no computational advantage against standard response formalism was found.[38] The SOS approach[29] used in this study is simpler using the approximate excited states derived from the Kohn–Sham determinant. Within the response-based TDDFT, for example, SOS sums involve a quadruple sum over the virtual and occupied molecular orbitals.[38] Note that the SOS and response approaches are equivalent only for exact wave functions and functionals.

Within the plain SOS approach, as far as we know, the GIAOs orbitals have not been used yet. Fortunately, when the origin is placed in the mass center of the molecule the error of computed MCD is small,[29] similarly as in ECD.[26] Nevertheless, it is highly desirable to remove the origin dependence also from the SOS methodology, because of the inherently unpredictable error it can cause, especially when more exotic or large systems are investigated.

In this work, we use an approach akin to the localized orbital/local origin method (LORG) previously developed[39] and used[40] for simulations of nuclear magnetic resonance (NMR) shielding. For MCD, the locality refers to the position of a center of charge of the investigated transition. With this choice of the coordinate origin, the ECD and MCD intensities appear to be fully origin-independent.

Below, we present the theory, including a formal proof that the LORG method provides fully origin-independent results. The implementation of the theory is tested on pyrrole and phenylalanine as typical examples studied by the circular dichroic spectroscopies.

Theory of MCD

The theory of MCD is well-documented in literature.[41–44] Traditionally, three molecular contributions to MCD intensities are recognized and referred to as the "Faraday's" *A*, *B*, and *C* terms. The *A*-terms can be expressed as a combination of *B*-terms.[13,29] As the *C*-term is not relevant to closed-shell molecules,[45] it is omitted in this study. For simplicity, we use the usual atomic units, where the electronic magnetic dipole moment, electric dipole moment, and gradient operators are defined as $\mathbf{m}=\frac{i}{2}\sum_{i=1}^{N_e}\mathbf{r}_i\times\nabla_i$, $\boldsymbol{\mu}=-\sum_{i=1}^{N_e}\mathbf{r}_i$, and $\nabla=\sum_{i=1}^{N_e}\nabla_i$, respectively, where $i=\sqrt{-1}$, N_e is the number of electrons, $\mathbf{r}_i$ is the position vector of *i*th electron, and $\nabla_i=\frac{\partial}{\partial\mathbf{r}_i}$. In these units, the dipole-velocity transformation valid for exact wave functions can be written as

$$\langle l|\boldsymbol{\mu}|n\rangle=\langle l|\nabla|n\rangle/(E_l-E_n) \tag{1}$$

where l and n are molecular electronic states, and E_l and E_n their energies. The index n is reserved for the electronic ground state, which we consider nondegenerate. For brevity, we also define $\boldsymbol{\mu}_{ln}=\langle l|\boldsymbol{\mu}|n\rangle$, $\nabla_{ln}=\langle l|\nabla|n\rangle$, and $E_{ln}=E_l-E_n$.

The canonical form of the *B*-term for a single $n\rightarrow j$ transition in an isotropic (rotationally averaged) sample is then[43]

$$B_r=\mathrm{Im}\left[\sum_{k\neq n}\frac{\mathbf{m}_{kn}\cdot\boldsymbol{\mu}_{nj}\times\boldsymbol{\mu}_{jk}}{E_{kn}}+\sum_{k\neq j}\frac{\mathbf{m}_{jk}\cdot\boldsymbol{\mu}_{nj}\times\boldsymbol{\mu}_{km}}{E_{kj}}\right] \tag{2}$$

We append the index "*r*" to the *B*-term to emphasize that the electric dipoles are written in the length form. Alternatively, using (1), we can rewrite the *B*-term as,

$$B_\nabla=\mathrm{Im}\left[\sum_{k\neq n,j}\left(\frac{\mathbf{m}_{kn}\cdot\nabla_{nj}\times\nabla_{jk}}{E_{kn}E_{nj}E_{jk}}+\frac{\mathbf{m}_{jk}\cdot\nabla_{nj}\times\nabla_{kn}}{E_{kj}E_{nj}E_{kn}}\right)+\frac{\mathbf{m}_{jn}\cdot\nabla_{nj}\times(\boldsymbol{\mu}_{nn}-\boldsymbol{\mu}_{jj})}{E_{nj}^2}\right] \tag{3}$$

which we refer to as the *B*-term expressed in the velocity (or gradient) formalism. B_∇ was found to be less origin-dependent than B_r.[29] However, when the molecule is placed far from the coordinate origin, the results based on B_∇ deteriorate very quickly as well. In contrast, "B_LORG" defined below will be origin-independent.

Origin-independent MCD SOS formulation

Unlike in the derivative ground-state response theory,[7–9] none of the SOS formulas (2, 3), thus removes the origin-dependence completely. However, this can be achieved by a LORG-like[39] compensating term. For this purpose, we rewrite eq. (2) again, this time as

$$\begin{aligned}B_\mathrm{LORG}=\frac{1}{2}\Bigg\{&\sum_{k\neq n}\left[\frac{\langle k|\sum_{i=1}^{N_e}\mathbf{r}_i\times\nabla_i|n\rangle}{E_{kn}}+\frac{1}{2N_e}\left(\sum_{l\neq n}\frac{\boldsymbol{\mu}_{kl}\times\nabla_{ln}}{E_{ln}}+\sum_{l\neq k}\frac{\boldsymbol{\mu}_{ln}\times\nabla_{kl}}{E_{k\mathrm{l}}}\right)\right]\cdot\boldsymbol{\mu}_{nj}\times\boldsymbol{\mu}_{jk}\\&+\sum_{k\neq j}\left[\frac{\langle j|\sum_{i=1}^{N_e}\mathbf{r}_i\times\nabla_i|k\rangle}{E_{kj}}+\frac{1}{2N_e}\left(\sum_{l\neq k}\frac{\boldsymbol{\mu}_{jl}\times\nabla_{lk}}{E_{kl}}+\sum_{l\neq j}\frac{\boldsymbol{\mu}_{lk}\times\nabla_{jl}}{E_{lj}}\right)\right]\cdot\boldsymbol{\mu}_{nj}\times\boldsymbol{\mu}_{kn}\Bigg\}\\=&\text{ ``}\mathbf{t1}+\mathbf{t2}+\mathbf{t3}+\mathbf{t4}+\mathbf{t5}+\mathbf{t6}\text{''}\end{aligned} \tag{4}$$

In the summations, we may sum over all l if we realize that the limit cases vanish. For the second term, for example, when $l=n$, we replace $(E_l-E_n)^{-1}$ by $(E_l-E_n)[(E_l-E_n)^2+\delta^2]^{-1}$, where δ is a small energy increment; then, because $\nabla_{nn}=0$, the contribution of **t2** is zero.

For exact wave functions due to the dipole-velocity transformation, terms **t2**, **t3**, **t5**, and **t6** are each zero as well. Taking again the second term as an example, and using the dipole-velocity transformation and the identities $\sum_l |l\rangle\langle l| = 1$ and $\boldsymbol{\mu}\times\boldsymbol{\mu}=0$, we obtain $\sum_l\sum_{k\neq n}\frac{\boldsymbol{\mu}_{kl}\times\nabla_{ln}}{N_e E_{ln}}\cdot\boldsymbol{\mu}_{nj}\times\boldsymbol{\mu}_{jk} = N_e{}^{-1}\sum_l\sum_{k\neq n}\boldsymbol{\mu}_{kl}\times\boldsymbol{\mu}_{ln}\cdot\boldsymbol{\mu}_{nj}\times\boldsymbol{\mu}_{jk} = N_e{}^{-1}\sum_{k\neq n}\langle k|\boldsymbol{\mu}\times\boldsymbol{\mu}|n\rangle\cdot\boldsymbol{\mu}_{nj}\times\boldsymbol{\mu}_{jk}=0.$

However, the B-term expressed in eq. (4) is origin-independent for any orthonormal set of wave functions. To demonstrate that, we just have to realize that for a coordinate origin shift by $\mathbf{T}$, $\mathbf{r}_i$ changes to $\mathbf{r}_i+\mathbf{T}$, $\boldsymbol{\mu}$ to $\boldsymbol{\mu}-N_e\mathbf{T}$ and B_{LORG} to $B_{\mathrm{LORG}}+\mathrm{X}$. Then

$$
\begin{aligned}
X=\frac{1}{2}\Bigg\{&\sum_{k\neq n}\left[\frac{\mathbf{T}\times\langle k|\sum_{i=1}^{N_e}\nabla_i|n\rangle}{E_{kn}}-\frac{N_e}{2N_e}\left(\sum_{l\neq n}\frac{\mathbf{T}\delta_{kl}\times\nabla_{ln}}{E_{ln}}+\sum_{l\neq k}\frac{\mathbf{T}\delta_{ln}\times\nabla_{kl}}{E_{kl}}\right)\right]\cdot\boldsymbol{\mu}_{nj}\times\boldsymbol{\mu}_{jk}\\
&+\sum_{k\neq j}\left[\frac{\mathbf{T}\times\langle j|\sum_{i=1}^{N_e}\nabla_i|k\rangle}{E_{kj}}-\frac{N_e}{2N_e}\left(\sum_{l\neq k}\frac{\mathbf{T}\delta_{jl}\times\nabla_{lk}}{E_{kl}}+\sum_{l\neq j}\frac{\mathbf{T}\delta_{lk}\times\nabla_{jl}}{E_{lj}}\right)\right]\cdot\boldsymbol{\mu}_{nj}\times\boldsymbol{\mu}_{kn}\\
&+\sum_{k\neq n}\left[\frac{\langle k|\sum_{i=1}^{N_e}\mathbf{r}_i\times\nabla_i|n\rangle}{E_{kn}}+\frac{1}{2N_e}\left(\sum_{l\neq n}\frac{\boldsymbol{\mu}_{kl}\times\nabla_{ln}}{E_{ln}}+\sum_{l\neq k}\frac{\boldsymbol{\mu}_{ln}\times\nabla_{kl}}{E_{kl}}\right)\right]\cdot\boldsymbol{\mu}_{nj}\times\mathbf{T}\delta_{jk}\\
&+\sum_{k\neq j}\left[\frac{\langle j|\sum_{i=1}^{N_e}\mathbf{r}_i\times\nabla_i|k\rangle}{E_{kj}}+\frac{1}{2N_e}\left(\sum_{l\neq k}\frac{\boldsymbol{\mu}_{jl}\times\nabla_{lk}}{E_{kl}}+\sum_{l\neq j}\frac{\boldsymbol{\mu}_{lk}\times\nabla_{jl}}{E_{lj}}\right)\right]\cdot\boldsymbol{\mu}_{nj}\times\mathbf{T}\delta_{kn}\Bigg\}\\
=\frac{1}{2}\Bigg\{&\sum_{k\neq n}\left[\frac{\mathbf{T}\times\nabla_{ln}}{E_{kn}}-\frac{1}{2}\left(\frac{\mathbf{T}\times\nabla_{ln}}{E_{kn}}+\frac{\mathbf{T}\times\nabla_{ln}}{E_{kn}}\right)\right]\cdot\boldsymbol{\mu}_{nj}\times\boldsymbol{\mu}_{jk}+\sum_{k\neq j}\left[\frac{\mathbf{T}\times\nabla_{jk}}{E_{kj}}-\frac{1}{2}\left(\frac{\mathbf{T}\times\nabla_{jk}}{E_{kj}}+\frac{\mathbf{T}\times\nabla_{jk}}{E_{kj}}\right)\right]\cdot\boldsymbol{\mu}_{nj}\times\boldsymbol{\mu}_{kn}\\
&+\left[\frac{\langle j|\sum_{i=1}^{N_e}\mathbf{r}_i\times\nabla_i|n\rangle}{E_{jn}}+\frac{1}{2N_e}\left(\sum_{l\neq n}\frac{\boldsymbol{\mu}_{jl}\times\nabla_{ln}}{E_{ln}}+\sum_{l\neq j}\frac{\boldsymbol{\mu}_{ln}\times\nabla_{jl}}{E_{jl}}\right)\right]\cdot\boldsymbol{\mu}_{nj}\times\mathbf{T}\\
&+\left[\frac{\langle j|\sum_{i=1}^{N_e}\mathbf{r}_i\times\nabla_i|n\rangle}{E_{nj}}+\frac{1}{2N_e}\left(\sum_{l\neq n}\frac{\boldsymbol{\mu}_{jl}\times\nabla_{ln}}{E_{nl}}+\sum_{l\neq j}\frac{\boldsymbol{\mu}_{ln}\times\nabla_{jl}}{E_{lj}}\right)\right]\cdot\boldsymbol{\mu}_{nj}\times\mathbf{T}\Bigg\}=0
\end{aligned}
$$

In fact, one can imagine many other "zero" terms added to eq. (2), which would compensate for the origin-dependence. The choice of eq. (4) is linked to the LORG formalism,[39] where the momentum operator is split as $\mathbf{l}=\mathbf{r}\times\mathbf{p}=(\mathbf{r}-\mathbf{r}_{\mathrm{loc}})\times\mathbf{p}+\mathbf{r}_{\mathrm{loc}}\times\mathbf{p}$. In formula (4), we can associate $\mathbf{r}_{\mathrm{loc}}=(\mathbf{r}_{kk}+\mathbf{r}_{nn})/2$ with each element of the magnetic integrals $\mathbf{m}_{kn}$ present in eq. (2), which has a clear physical meaning as an average center of charge for the $k\rightarrow n$ transition. The additional summations in eq. (4), if compared with eqs. (2) and (3), do not lead to a significant increase of computational time for our systems, as this is primarily limited by the TDDFT computations of the excited molecular states.

Origin-independent formulation of ECD

As described elsewhere,[26,44,46] the rotational strength determining ECD spectral intensities for a transition $n\rightarrow j$ can be calculated within the length and gradient-based formalisms as

$$R_r=\frac{1}{2}\langle n|\sum_{i=1}^{N_e}\mathbf{r}_i\times\nabla_i|j\rangle\cdot\boldsymbol{\mu}_{jn}\qquad(5)$$

$$R_\nabla=\frac{1}{2E_{jn}}\langle n|\sum_{i=1}^{N_e}\mathbf{r}_i\times\nabla_i|j\rangle\cdot\nabla_{jn}\qquad(6)$$

The first formulation [eq. (5)] is origin-dependent. However, when the molecule is close to the origin, eq. (5) is generally considered more accurate as the gradient operator (∇_{jn}) is absent.[46] Formula (6) is "safer to use" as it is origin-independent, but the gradient operator makes it more sensitive to the quality of the wave function.

Fortunately, with a similar reasoning as for MCD, the ECD rotational strength can be expressed as

$$R_{\mathrm{LORG}}=\frac{1}{2}\left(\langle n|\sum_{i=1}^{N_e}\mathbf{r}_i\times\nabla_i|j\rangle-\left(\frac{\boldsymbol{\mu}_{nn}+\boldsymbol{\mu}_{jj}}{2N_e}\right)\times\nabla_{nj}\right)\cdot\boldsymbol{\mu}_{jn},\qquad(7)$$

that is, as an origin-independent length-based formula combining the advantages of eqs. (5) and (6). The additional second term in eq. (7), zero for exact wave functions [when $\nabla_{nj}\sim\boldsymbol{\mu}_{jn}$, see eq. (1)], is compensating for the origin-dependence of the first one.

Ab initio computations

The test molecules included pyrrole previously used in MCD computational tests[29] and phenylalanine as a typical molecule

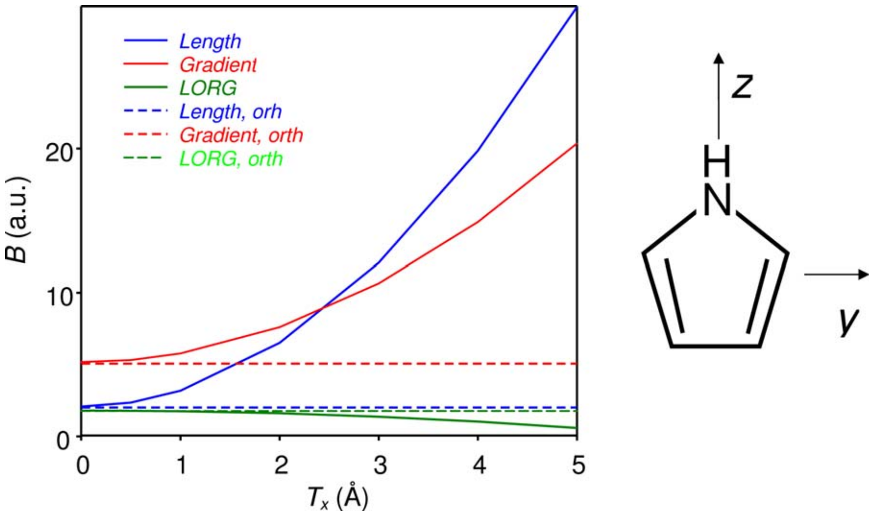

Figure 1. Pyrrole: dependence of the *B*-term for the 203.7 nm transition on the distance from coordinate origin. The length, gradient, and LORG forms of *B* are estimated without and with the additional orthonormalization (orth) of the excited states.

encountered in MCD and ECD studies. The Gaussian program (v. D01)[47] was used for the geometry optimizations and TDDFT computations. The spectral intensities were calculated with our own software[29] interfaced to Gaussian. Standard basis sets and the B3LYP[48] density functional theory functional were used as specified later. All single-excited states were used to calculate the spectra (except for pyrrole and the aug-cc-pVQZ basis set, where first 1000 states were calculated). The MCD, ECD, and absorption spectra were simulated using Gaussian bands of 10 nm full width at half height.

Results and Discussion

Origin-dependence versus state normalization

We found it very important to renormalize and reorthogonalize the TDDFT electronic states as these were obtained with a limited accuracy from the output of the quantum-chemical program. Owing to the symmetry of the pyrrole molecule, we studied earlier,[29] all *B*-term expressions (B_r, B_∇, and B_{LORG}) become origin-independent only when the states satisfy a condition $\langle i|j\rangle = \delta_{ij}$. This is shown in Figure 1 for the strong 203.7 nm transition calculated at the B3LYP/6-311++G** level of approximation. The *B*-terms are calculated with the coordinate origin shifted from the center of mass by 0–5 Å in the *x*-direction. The solid lines connect points obtained by the raw output, that is, single-excited state coefficients were cut at the limit of 10^{-4}, and read up to the fifth digit beyond the floating point. The dashed lines are used for coefficients renormalized by the Gram–Schmidt orthogonalization, repeated twice for a better numerical accuracy.

Clearly, the relatively small inaccuracy of the excited states (the default coefficient cutoff of 10^{-4} corresponds to the probability of mere 10^{-8}!) has serious implications for the origin-dependence. As discussed earlier,[29] the gradient form B_∇ is less dependent on the origin shift than B_r. These two approaches in general provide different results even for a zero shift, due to the incomplete basis set and inaccurate TDDFT electronic state energies. Note that for the zero shift, B_r is supposedly more accurate than B_∇, as the additional gradient operators introduce errors with incomplete basis functions.[46] The LORG approach appears to be as the least susceptible to the origin shift even if the wave function is inaccurate. Indeed, the B_r and B_{LORG} values almost coincide for the zero shift.

Although the renormalization is easy to do (and from this point, it is used throughout this study), the susceptibility to the TDDFT wave function error reveals important differences in the convergence behavior of the different approaches to MCD simulation.

Basis set convergence

For the pyrrole molecule with the coordinate origin placed in the center of mass, the calculated dependence of B_r, B_∇, and B_{LORG} on the number of atomic orbitals (6-31G, 6-31G**, 6-311++G**, 6-311++G**(2df), aug-cc-pVTZ, and aug-cc-pVQZ basis sets) is plotted in Figure 2 for the 203.7 nm (B3LYP/6-311++G**) transition. Other transitions behaved similarly. Clearly, all approaches exhibit a strong dependence on the size of the basis set, in accord with previous observations for MCD computations.[8,16,45] For the two smallest basis sets, 6-31G and 6-31G**, MCD is even reproduced with a wrong sign.

For the other larger basis sets containing both the polarization and diffuse functions, the results start to be more consistent and converging. Only a relatively very small change in the *B*-term is due to the aug-cc-pVTZ → aug-cc-pVQZ basis set upgrade. This is important for practical computations, as this upgrade translates to a significant increase of computational time (3.5 h vs. 6.5 days on our computer).

The earlier described sensitivity to the choice of the basis set of the B_r and B_∇ values reflect the inaccuracy of the TDDFT electronic excited states, that is, primarily the error of the B3LYP functional. The B_{LORG} values closely match the B_r results; however, note that this occurs only for symmetric molecules or coordinate center close to the center of mass.

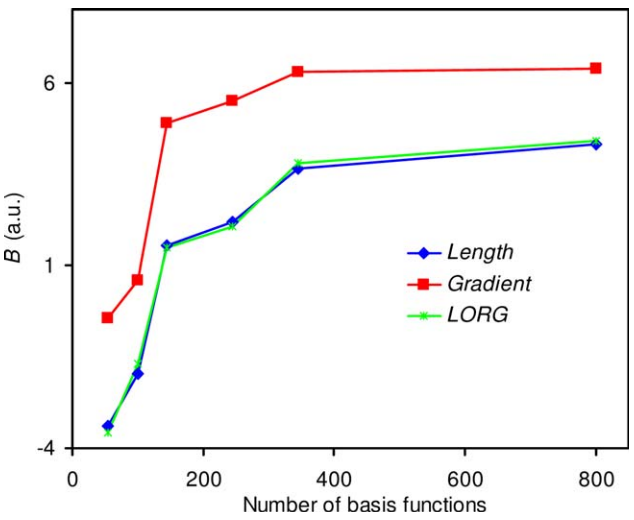

Figure 2. Pyrrole; dependence of the *B*-term for the 203.7 nm transition on the number of atomic orbitals as calculated by the three approaches. (Origin coincided with the center of mass, for B3LYP/6-311+G** optimized geometry). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Origin-dependence: ECD and MCD of phenylalanine

The phenylalanine zwitterion has a relatively strong ECD and MCD signal, which needs to be taken into account for structural studies of proteins if using these spectroscopic techniques.[49,50] The molecule has a lower symmetry (C_1) than pyrrole (C_{2v}). Therefore, the calculated MCD spectra remain origin-dependent even after the state orthogonalization, unless the LORG formalism is used. This is shown in Figure 3 for three origin shifts (0, 2, and 10 Å from the center of mass).

For 0 Å, the length-based and LORG results almost coincide. The spectral curve obtained by the gradient formalism is slightly different. For example, the positive signal at ~180 nm present in the "gradient" spectrum is visible in the length formalism as a local maximum only. The shift of 2 Å causes relatively minor changes in the B_r curve, whereas for 10 Å, the spectrum simulated by the length formalism is far from satisfactory.

The gradient approach is thus significantly less dependent on the origin shift than the length one. Nevertheless, both of them exhibit significant differences compared with the reference computation when the origin is placed at the center of mass. Only the LORG approach provides truly independent SOS simulation of MCD.

The ECD spectra of phenylalanine obtained within the length, gradient, and LORG approach are plotted in Figure 4, as calculated with the origin at the mass center, and shifted in the *x*-direction by 10 and 25 Å. Even the length approach provides spectra that are significantly more resistant to the origin shift than for MCD. For example, the origin shift of 10 Å causes ECD intensity variations up to about 50%, that is, the spectral shape is deformed much less than for MCD (cf. Fig. 3). Nevertheless, as expected, only the gradient and LORG methods provide origin-independent ECD curves. For the reference zero shift, the gradient ECD intensity is within 5% identical to that obtained by the length approach. Therefore, at this level of approximation, the gradient-based approach provides origin-independent simulations that are accurate enough for most applications. Yet the LORG simulation

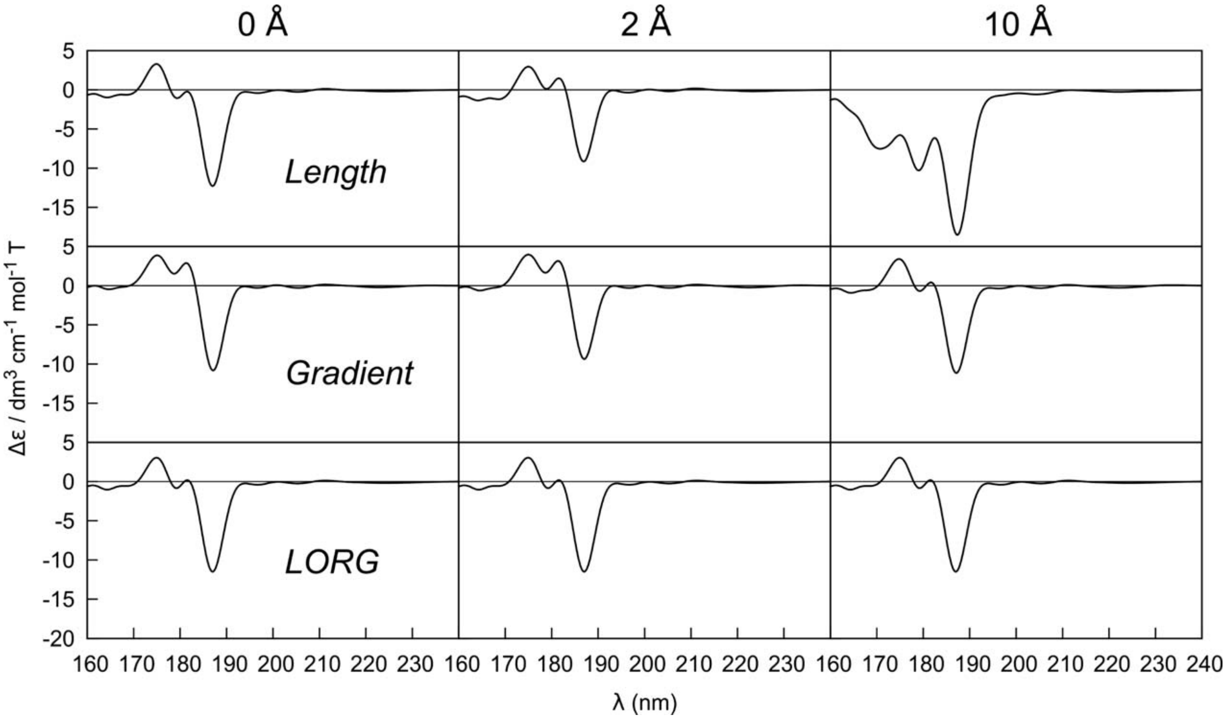

Figure 3. Phenylalanine, MCD spectra simulated by the three SOS approaches for the coordinate origin placed at 0, 2, and 10 Å far from molecular center of mass. Note that unlike for pyrrole (Fig. 1), the length and gradient methods are origin-dependent even with the additional orthonormalization.

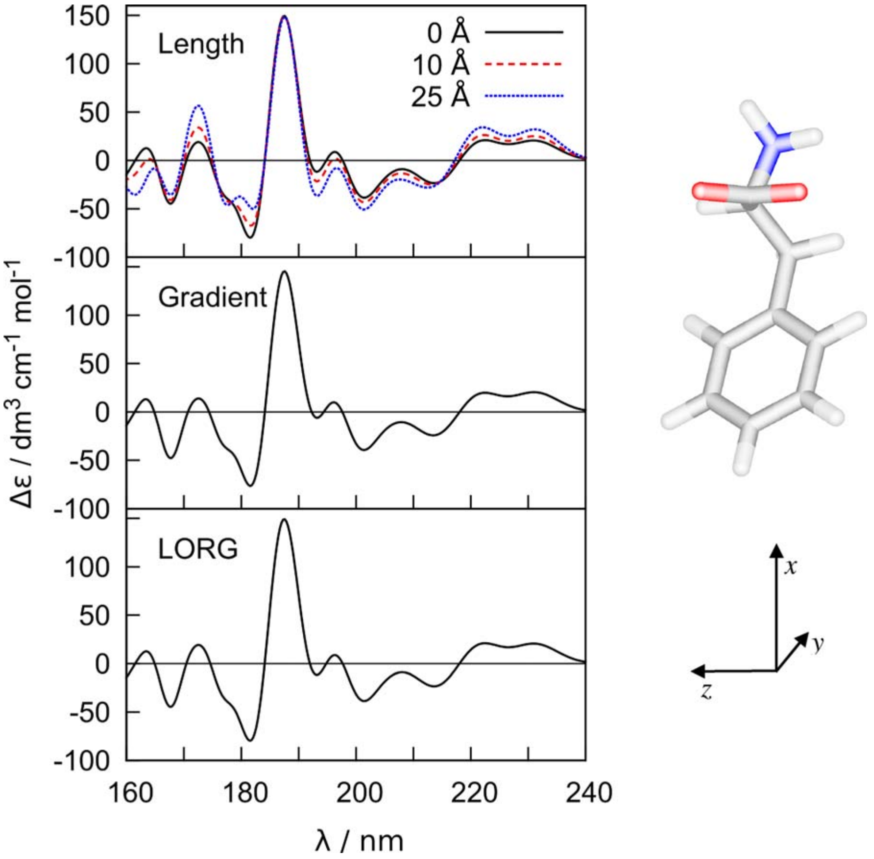

Figure 4. Phenylalanine, ECD spectra simulated by the length, gradient, and LORG approaches for the coordinate origin shifted byt 0, 10, and 25 Å away from the center of mass, in the *x*-direction. Note that the 0, 10, and 25 Å curves are identical in the gradient and LORG approaches.

provides intensities even closer to the reference simulation, within 0.1% identical to those obtained by the length formalism and zero origin shift.

The dependence of ECD on the basis set is documented on the rotational strength calculated for the second lowest-energy phenylalanine transition in Figure 5. The molecular mass center was *x*-shifted by 25 Å from the origin. Apparently, the length-based rotational strength converges slightly less smoothly and significantly deviates from the values obtained by the LORG or the gradient. The length-based method thus combines errors stemming from both the incomplete basis set and the inherent origin-dependence. As expected, R_r and R_{LORG} approach each other for the largest aug-cc-pVTZ basis set.

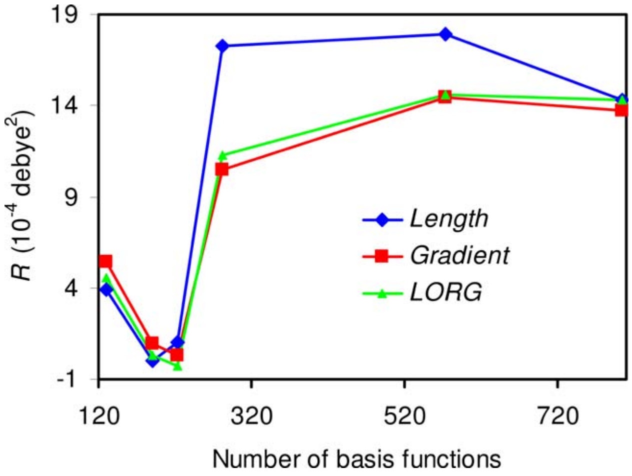

Figure 5. Phenylalanine, dependence of the rotational strength on the number of basis set functions, as obtained by the three methods (for 6-31G, 6-31G*, 6-31G**, 6-31+G*, 6-311++G(2df,2pd), and aug-cc-pVTZ basis sets, B3LYP functional, mass center 25 Å away from the origin, for a transition calculated at 232 nm with the aug-cc-pVTZ basis set). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Unlike for MCD (Fig. 2), the gradient and length ECD formalisms (Fig. 5) appear to approach a common infinite basis set limit. MCD involving transitions between excited states is in this respect more strongly affected by approximating the true wave function by the states based on the Kohn–Sham orbitals. This may be due to the adiabatic and other approximations used in the exchange-correlation potential. Indeed, MCD *B*-terms based on the states derived from the time-dependent Hartree–Fock (TDHF) theory exhibit a better convergence of the gradient and length values, with respect to the number of basis functions than TDDFT. This is shown in Figure 6, where the SOS *B*-term of the lowest-energy transition in the phenylalanine model molecule is shown as calculated with several basis sets. At the TDHF level, all the length, gradient, and LORG SOS approaches appear to converge to a similar value. Conversely, for TDDFT with the B3LYP functional, the gradient approximation yields a different limit than the other length-based formalisms, which is consistent with the pyrrole data (cf. Fig. 2).

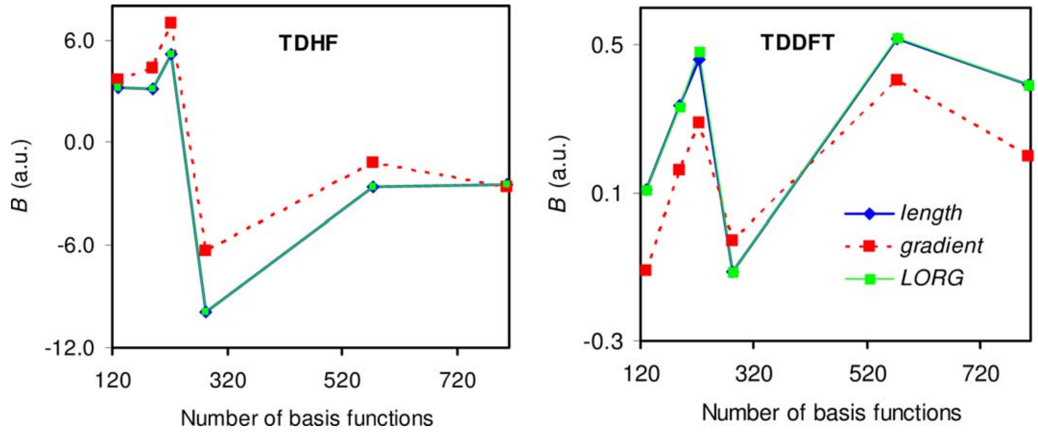

Figure 6. MCD *B*-term of phenylalanine computed using the three SOS approaches and TDHF (left) and TDDFT (right) states: dependence on the number of basis functions. The basis sets were same as in Figure 5, for the lowest-energy transition calculated with the aug-cc-pVTZ basis set at 213 and 234 nm. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Conclusions

This study investigated the origin-dependence of the pure sum over state formalism of generating MCD spectra. For the MCD *B*-term, we derived a LORG expression with origin-compensating terms and proved its origin-independence. Pyrrole and phenylalanine zwitterion were used as test systems. Unlike the length and gradient forms, the LORG procedure provided origin-independent MCD intensities. Thus it is possible to use more robust length-computations of the transition electric dipole moments, which converges reasonably well with the size of the basis set.

A similar LORG-like correction was added to the length-based formula for the rotational strength, which made the simulated ECD spectra origin-independent. We also showed that ECD is more “immune” to shifting of the coordinate origin than MCD. The length-based formula thus might be usable in applied computations if molecular center of mass is close to the origin of coordinates. Generally, however, the LORG-correction appears more advantageous yet does not bring about any significant computational burden compared with the conventional procedures. The results can thus be immediately used in molecular ECD and MCD simulations.

Keywords: density functional theory · electronic circular dichroism · magnetic circular dichroism · origin-dependence

How to cite this article: P. Štěpánek, P. Bouř *J. Comput. Chem.* **2015**, *36*, 723–730. DOI: 10.1002/jcc.23845

[1] N. Berova, K. Nakanishi, R. W. Woody, Eds. Circular Dichroism Principles and Applications; Wiley-VCH: New York, **2000**.

[2] T. Xiang, D. J. Goss, M. Diem, *Biophys. J.* **1993**, *65*, 1255.

[3] W. C. Johnson, Molecular Chirality in Chemistry and Life Sciences; CNR: Pisa, Italy, **1997**; p. 15.

[4] W. C. Johnson, In Circular Dichroism and the Conformational Analysis of Biomolecules; G. D. Fasman, Ed.; Plenum Press: New York, **1996**; p. 433.

[5] C. Angeli, K. L. Bak, V. Bakken, O. Christiansen, R. Cimiraglia, S. Coriani, P. Dahle, E. K. Dalskov, T. Enevoldsen, B. Fernandez, C. Haettig, K. Hald, A. Halkier, H. Heiberg, T. Helgaker, H. Hettema, H. J. A. Jensen, D. Jonsson, P. Joergensen, S. Kirpekar, W. Klopper, R. Kobayashi, H. Koch, O. B. Lutnaes, K. V. Mikkelsen, P. Norman, J. Olsen, M. J. Packer, T. B. Pedersen, Z. Rinkevicius, E. Rudberg, T. A. Ruden, K. Ruud, P. Salek, A. Sanchez de Meras, T. Saue, S. P. A. Sauer, B. Schimmelpfennig, K. O. Sylvester-Hvid, P. R. Taylor, O. Vahtras, D. J. Wilson, H. Agren, Dalton 2013, University of Oslo: Oslo, **2013**.

[6] E. J. Baerends, T. Ziegler, J. Autschbach, D. Bashford, A. Bérces, F. M. Bickelhaupt, C. Bo, P. M. Boerrigter, L. Cavallo, D. P. Chong, L. Deng, R. M. Dickson, D. E. Ellis, M. van Faassen, L. Fan, T. H. Fischer, C. F. Guerra, M. Franchini, A. Ghysels, A. Giammona, S. J. A. van Gisbergen, A. W. Götz, J. A. Groeneveld, O. V. Gritsenko, M. Grüning, S. Gusarov, F. E. Harris, P. van den Hoek, C. R. Jacob, H. Jacobsen, L. Jensen, J. W. Kaminski, G. van Kessel, F. Kootstra, A. Kovalenko, M. V. Krykunov, E. van Lenthe, D. A. McCormack, A. Michalak, M. Mitoraj, S. M. Morton, J. Neugebauer, V. P. Nicu, L. Noodleman, V. P. Osinga, S. Patchkovskii, M. Pavanello, P. H. T. Philipsen, D. Post, C. C. Pye, W. Ravenek, J. I. Rodríguez, P. Ros, P. R. T. Schipper, G. Schreckenbach, J. S. Seldenthuis, M. Seth, J. G. Snijders, M. Solà, M. Swart, D. Swerhone, G. te Velde, P. Vernooijs, L. Versluis, L. Visscher, O. Visser, F. Wang, T. A. Wesolowski, E. M. van Wezenbeek, G. Wiesenekker, S. K. Wolff, T. K. Woo, A. L. Yakovlev, ADF modeling suite 2013, Theoretical Chemistry; Vrije Universiteit: Amsterdam, The Netherlands, **2013**.

[7] M. Seth, J. Autschbach, T. Ziegler, *J. Chem. Theory Comput*. **2007**, *3*, 434.

[8] T. Kjaergaard, S. Coriani, K. Ruud, *Wiley Interdiscip. Rev. Comput. Mol. Sci*. **2012**, *2*, 443.

[9] H. Solheim, K. Ruud, S. Coriani, P. Norman, *J. Chem. Phys*. **2008**, *128*, 094103.

[10] H. Solheim, K. Ruud, S. Coriani, P. Norman, *J. Phys. Chem. A* **2008**, *112*, 9615.

[11] P. Štěpánek, V. Andrushchenko, K. Ruud, P. Bouř, *J. Phys. Chem. A* **2012**, *116*, 778.

[12] G. A. Peralta, M. Seth, H. Zhekova, T. Ziegler, *Inorg. Chem*. **2008**, *47*, 4185.

[13] W. R. Mason, A Practical Guide to Magnetic Circular Dichroism Spectroscopy; Wiley-Interscience: Portland, **2007**.

[14] J. Mack, M. J. Stillman, N. Kobayashi, *Coord. Chem. Rev*. **2007**, *251*, 429.

[15] J. Michl, *Tetrahedron* **1984**, *40*, 3845.

[16] P. Štěpánek, M. Straka, V. Andrushchenko, P. Bouř, *J. Chem. Phys*. **2013**, *138*, 151103.

[17] H. Solheim, K. Kornobis, K. Ruud, P. M. Kozlowski, *J. Phys. Chem. B* **2011**, *115*, 737.

[18] J. Autschbach, *Chem. Phys. Chem*. **2011**, *12*, 3224.

[19] M. Krykunov, J. Autschbach, *J. Chem. Phys*. **2005**, *123*, 114103.

[20] K. L. Bak, A. E. Hansen, K. Ruud, T. Helgaker, J. Olsen, P. Jørgensen, *Theor. Chim. Acta*. **1995**, *90*, 441.

[21] M. Pecul, K. Ruud, T. Helgaker, *Chem. Phys. Lett*. **2004**, *388*, 110.

[22] S. Coriani, C. Hättig, P. Jørgensen, *J. Chem. Phys*. **2000**, *113*, 3561.

[23] M. Seth, T. Ziegler, *J. Chem. Phys*. **2007**, *127*, 134108.

[24] M. Krykunov, M. Seth, T. Ziegler, J. Autschbach, *J. Chem. Phys*. **2007**, *127*, 244102.

[25] K. M. Lee, K. Yabana, G. F. Bertsch, *J. Chem. Phys*. **2011**, *134*, 144106.

[26] J. Autschbach, T. Ziegler, S. J. A. van Gisbergen, E. J. Baerends, *J. Chem. Phys*. **2002**, *116*, 6930.

[27] F. Mauri, B. G. Pfrommer, S. G. Louie, *Phys. Rev. Lett*. **1996**, *77*, 5300.

[28] C. J. Pickard, F. Mauri, *Phys. Rev. B* **2001**, *6324*, 245101.

[29] P. Štěpánek, P. Bouř, *J. Comput. Chem*. **2013**, *34*, 1531.

[30] P. Štěpánek, P. Bouř, *Phys. Chem. Chem. Phys*. 2014, 16, 20639.

[31] V. Andrushchenko, D. Padula, E. Zhivotova, S. Yamamoto, P. Bouř, *Chirality* **2014**, *26*, 655.

[32] R. G. Parr, W. Yang, Density-Functional Theory of Atoms and Molecules; Oxford University Press: New York, **1994**.

[33] P. Bouř, *J. Comput. Chem*. **2000**, *21*, 8.

[34] P. Bouř, *J. Phys. Chem. A* **1999**, *103*, 5099.

[35] D. W. Miles, H. Eyring, *Proc. Natl. Acad. Sci. USA* **1973**, *70*, 3754.

[36] E. Goldstein, S. Vijaya, G. A. Segal, *J. Am. Chem. Soc*. **1980**, *102*, 6198.

[37] A. R. Meier, H. Wagniere, *Chem. Phys*. **1987**, *113*, 287.

[38] J. Autschbach, M. Seth, T. Ziegler, *J. Chem. Phys*. **2007**, *126*, 174103.

[39] A. E. Hansen, T. D. Bouman, *J. Chem. Phys*. **1985**, *82*, 5035.

[40] P. Bouř, *Int. J. Quantum Chem*. **2003**, *91*, 277.

[41] R. Serber, *Phys. Rev*. **1932**, *41*, 489.

[42] P. J. Stephens, *J. Chem. Phys*. **1970**, *52*, 3489.

[43] P. J. Stephens, *Annu. Rev. Phys. Chem*. **1974**, *25*, 201.

[44] L. D. Barron, Molecular Light Scattering and Optical Activity; Cambridge University Press: Cambridge, **2004**.

[45] M. Seth, T. Ziegler, *J. Chem. Phys*. **2006**, *124*, 144105.

[46] S. Grimme, F. Furche, R. Ahlrichs, *Chem. Phys. Lett*. **2002**, *361*, 321.

[47] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian 09, Revision D01 Inc.: Wallingford CT, **2014**.

[48] A. D. Becke, *J. Chem. Phys*. **1993**, *98*, 5648.

[49] C. Djerassi, G. Barth, R. Records, E. Bunnenberg, W. Voelter, *J. Am. Chem. Soc*. **1971**, *93*, 2545.

[50] N. Sreerama, M. C. Manning, M. E. Powers, J. X. Zhang, D. P. Goldenberg, R. W. Woody, *Biochemistry* **1999**, *38*, 10814.

Received: 27 November 2014
Revised: 5 January 2015
Accepted: 9 January 2015