Journal of Physics: Conference Series **698** (2016) 012009 doi:10.1088/1742-6596/698/1/012009

# Can We Describe Biological Systems with Quantum Mechanics?

**C G Granados-Ramírez[1], C G Benítez-Cardoza[2], and M D Carbajal-Tinoco[3]**

[1] Grupo de Investigación Bioquímica INS & LIBBIQ-UN Universidad Nacional de Colombia, Cra. 45 No. 26-85 U. Camilo Torres, Bogotá 111321, Colombia
[2] Laboratorio de Investigación Bioquímica, Sección de Estudios de Posgrado, ENMyH-IPN, Guillermo Massieu No. 239, La Escalera Ticomán, 07320 México D.F., Mexico
[3] Departamento de Física, Centro de Investigación y de Estudios Avanzados del IPN, Apartado Postal 14-740, 07000 México D.F., Mexico

E-mail: mdct@fis.cinvestav.mx

**Abstract.** Quantum Mechanics is the favourite theory to predict the structure of any group of atoms, including biological molecules. Due to numerous difficulties, however, it is necessary to introduce a series of approximations to overcome such impediments. We present a coarse-grained model of circular dichroism (CD) that is based on the theory of optical activity, developed by DeVoe, in order to predict CD spectra. In first stage, we determine the polarisability of individual monomers (residues, in the case of peptides) from experiments of molar absorptivity. The complex polarisabilities are used together with peptide structures obtained by density functional theory and other methods to determine their corresponding CD spectra, which are in reasonable agreement with their experimental counterparts.

## 1. Introduction

The description and full understanding of biological systems is one of the major scientific tasks for this century. For example, malaria killed about half a million people around the world in 2015 and that is good news because it used to kill twice as many persons, 20 years earlier. In 2015, Youyou Tu received the Nobel prize in medicine for using a plant extract, artemisinin, in the treatment of malaria. In spite of important efforts, there is still no vaccine for malaria and the complete cure remains an open problem as in the case of many other diseases. The comprehension of biological systems can be a challenge because they are composed of a variety of molecules such as lipids, proteins, and nucleic acids like DNA and RNA. These molecules have specific functions that include interaction with other molecules and, eventually, they can modify their shape to perform a certain activity.

The typical biological systems are characterised by experimentally accessible distances (from nanometers to microns) and relatively low energies (of the order of the thermal energy, $k_BT$). Therefore, the preferred approach to describe such systems should be Quantum Mechanics (QM). For instance, we can propose the solution of the time-independent non-relativistic Schrödinger equations for a system consisting of $M$ nuclei and $N$ electrons that mainly interact through a Coulomb potential. A reduced description, that captures the intended behaviour of the system, requires more than $10^6$ atoms (including water molecules). Thus, it is basically impossible

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

to solve the resulting highly-coupled Schrödinger equations, even at the level of the Born-Oppenheimer approximation. The exclusion of the solvent molecules provides a significant reduction of the number of atoms, but it is still necessary to utilise an appropriate theoretical framework. For this purpose, we briefly describe the Density Functional Theory (DFT) in section 2.

In this work, we used DFT to obtain the two stable configurations of a relatively small peptide, namely, polylysine of 150 residues. Such structures are then utilised as a reference in a proposed model of circular dichroism (CD), which is an experimental technique that takes advantage of the chiral nature of biological molecules, allowing the identification of structural changes. Once calibrated, we make use of additional approximations in order to compare the results of our model with experimental data, as explained in the following sections.

## 2. Theoretical background

### *2.1. Density functional theory*

The DFT is a computational method that allows to perform QM modelling of divers systems including condensed-matter physics, quantum chemistry, and the ground-state of a variety of molecules that could be of biological origin. A good review on the subject can be consulted in Ref. [1]. The electron density $\rho(\vec{r})$ is the central quantity in DFT, since it determines the probability of finding any of the $N$ electrons of the system within the volume element $d\vec{r}$. It is defined as the integral over the spin coordinates of all electrons,

$$\rho(\vec{r}_1) = N \int \cdots \int |\Psi(\vec{x}_1, \vec{x}_2, ..., \vec{x}_N)|^2 ds_1 d\vec{x}_1 \cdots d\vec{x}_N, \tag{1}$$

together with the normalisation condition, $\int \rho(\vec{r}) d\vec{r} = N$. The electron density is a positive function that can be experimentally measured and it shows an exponential decay for large distances, $\rho(\vec{r}) \sim \exp[2\sqrt{2I}|\vec{r}|]$, where $I$ is the exact ionization energy. The determination of $\rho(\vec{r})$ is based on two fundamental theorems established by Hohenberg and Kohn (H-K) in 1964 [2]. For an external potential $v(\vec{r})$ acting on the $N$ electrons of the system, the first H-K theorem states that $v(\vec{r})$ is determined (to within a constant) by the electron density $\rho(\vec{r})$. In other words, $v(\vec{r})$ is a unique functional of $\rho(\vec{r})$. The second H-K theorem states that for all $v$-representable densities, $E_v[\rho] \geq E_v[\rho_0]$, with $E_v[\rho_0]$ being the ground-state energy and $\rho_0$ is the corresponding ground-state density. The second theorem is of course a direct consequence of the variational principle [3].

Besides the H-K theorems, the Kohn-Sham (K-S) equations can be determined as a function of $N$ orbitals $\psi_i$ [4]. Starting with the expression for the energy functional $E[\rho]$ that depends on the interactions between the $N$ electrons and $M$ nuclei of the system,

$$\begin{aligned} E[\rho] &= -\frac{1}{2}\sum_i^N \langle\psi_i|\nabla^2|\psi_i\rangle + \frac{1}{2}\sum_i^N\sum_j^N \int\int |\psi_i(\vec{r}_1)|^2 \frac{1}{r_{12}} |\psi_j(\vec{r}_2)|^2 d\vec{r}_1 d\vec{r}_2 \\ &\quad + E_{XC}[\rho] - \sum_i^N \int \sum_A^M \frac{Z_A}{r_{1A}} |\psi_i(\vec{r}_1)|^2 d\vec{r}_1, \end{aligned} \tag{2}$$

with $Z_A$ being the valence of nucleus $A$, $r_{12} = |\vec{r}_2 - \vec{r}_1|$, and the explicit form of the term $E_{XC}[\rho]$ is unknown in general. Applying the usual constraint $\langle\psi_i|\psi_j\rangle = \delta_{ij}$, leads to the set of equations,

$$\left(-\frac{1}{2}\nabla^2 + V_S(\vec{r}_1)\right)\psi_i = \epsilon_i \psi_i, \tag{3}$$

where,

$$V_S(\vec{r}_1) = \int \frac{\rho(\vec{r}_2)}{r_{12}} d\vec{r}_2 + V_{XC}(\vec{r}_1) - \sum_A^M \frac{Z_A}{r_{1A}},$$

$\epsilon_i$ are the energy eigenvalues of orbitals $\psi_i$, and $V_{XC} = \delta E_{XC}/\delta\rho$. At this point, it is necessary to provide an approximation for $E_{XC}$. In the generalised gradient approximation (GGA) there is a dependence on densities and their gradients,

$$E_{XC}^{GGA}[\rho_\alpha, \rho_\beta] = \int f(\rho_\alpha, \rho_\beta, \nabla\rho_\alpha, \nabla\rho_\beta) d\vec{r}. \tag{4}$$

This term can be combined with the K-S contribution to obtain a hybrid functional, namely,

$$E_{XC}^{hyb} = aE_{XC}^{KS} + (1-a)E_{XC}^{GGA}, \tag{5}$$

here, $a$ is a fitting parameter. In the K-S scheme, the orbitals are expanded in a set of $L$ predefined basis functions $\eta_\mu$,

$$\psi_i = \sum_{\mu=1}^{L} c_{\mu i}\eta_\mu. \tag{6}$$

The Gaussian-type-orbitals are a convenient choice for the basis functions, i.e., $\eta^l \sim r^l \exp[-\kappa r^2]$.

We obtained the ground-state of a poly-L-lysine (also denoted as polylysine, for short) of 150 residues in vacuum by using the package Orca [5] and the density functional B3LYP [6]. This peptide has two stable configurations, namely, $\alpha$-helix and $\beta$-sheet, which are employed to determine two constants parameters of our CD model. On the other hand, modelling a polylysine in aqueous solution is nowadays out of the scope of DFT not only from the computational point of view. There are still some difficulties in describing certain kinds of intermolecular interactions such as dispersion forces. Instead, we made use of a classical force-field to model our peptides dissolved in water.

### *2.2. Circular dichroism*

An important feature of biological molecules is that a considerable number of them are chiral, in other words, they are not identical to their mirror images. Chiral structures can be distinguished and characterised by polarised light [7]. CD is the differential absorption of left and right circularly polarised light or the difference between their respective extinction coefficients,

$$\Delta\varepsilon = \varepsilon_L - \varepsilon_R. \tag{7}$$

The rotational strength of a transition $A \leftarrow 0$ is the integrated intensity of a single wavelength band in the CD spectrum

$$R_{0A} \sim \int (\Delta\varepsilon/\lambda) d\lambda, \qquad R_{0A} = Im(\langle\Psi_0|\vec{\mu}_e|\Psi_A\rangle \cdot \langle\Psi_0|\vec{\mu}_m|\Psi_A\rangle), \tag{8}$$

with $\vec{\mu}_e$ and $\vec{\mu}_m$ being the electronic and magnetic transition dipole moments. There are theoretical calculations of CD spectra within this QM formalism that are rather involved, but they are only in qualitative agreement with the corresponding experimental data [8] and that is a consequence of the used approximations. In this paper, we present an approach that combines the accuracy of the structures obtained by means of DTF together with experimental

data and the results of all-atom (classical) molecular dynamics (MD). The resulting model is computationally tractable even for proteins of hundreds of residues.

Fifty years ago, DeVoe proposed a classical theory of CD [9]. For an applied electric field $\vec{\mathbf{E}}_i$, the induced dipole moment $\vec{\mu}_i$ at each residue of polarisability tensor $\hat{\alpha}$ and components $\alpha_i$, is given by

$$\vec{\mu}_i = \alpha_i \left[ \vec{\mathbf{E}}_i + \sum_{i \neq j} T_{ij} \vec{\mu}_i \right], \tag{9}$$

where the components of the dipole interaction tensor $\hat{T}$ are defined as,

$$T_{ij} = -\left[ 3p_i q_j - \delta_{pq} r_{ij}^2 \right] / r_{ij}^5, \tag{10}$$

and $(p, q = x, y, z)$. The formal solution of Eq. (9) is $\hat{\mu} = \hat{B}\hat{\mathbf{E}}$ with $\vec{\mu}_i = |\vec{\mu}_i|\vec{\mathbf{e}}_i$. It can be assumed that an average polarisability $\alpha$ is able to approximate the three principal components of the monomer polarisabilities [10]. An advantage of the polarisability $\alpha$ is that it can be extracted from experimental data, as explained in the following lines. According to the Beer-Lambert law, the monomer extinction coefficient $\varepsilon$ is proportional to the absorbance [7], which is a measurable quantity. Additionally, we use a first order approximation to relate the extinction coefficient with the imaginary part of $\alpha$, i.e.,

$$\varepsilon \simeq \frac{k_1}{\lambda} \sum_{i=1}^{m} Im(\alpha_i). \tag{11}$$

Here, $k_1$ is a constant, $m$ is the total number of identical monomers (residues, in our case), and $\alpha_i = \alpha(\omega)$ with $\omega$ being the angular frequency. The real part of $\alpha(\omega)$ is obtained through a Kramers-Kronig transform,

$$Re(\alpha(\omega)) = -\frac{2}{\pi} \int_0^\infty \frac{Im(\alpha(\omega'))\omega' d\omega'}{[\omega'^2 - \omega^2]}. \tag{12}$$

On the other hand, CD can be determined through the ellipticity $[\theta] \sim \Delta\varepsilon$,

$$[\theta] = \frac{k_2}{\lambda^2} \sum_{i,j=1}^{m} (\vec{\mathbf{e}}_i \times \vec{\mathbf{e}}_j) \cdot (\vec{r}_j - \vec{r}_i) A_{ij}, \tag{13}$$

where $\hat{A} = [\hat{\alpha}^{-1} + \hat{G}]^{-1}$, $G_{ij} = T_{ij}\vec{\mathbf{e}}_i \cdot \vec{\mathbf{e}}_j$, and the constant $k_2$ has to be determined. The results of this model are presented in the next section.

## 3. Experimental part

In this section, we first depict the synthesis of our peptides and the experimental techniques used to describe them. These results together with the DFT and MD calculations are the main ingredients of our model of CD.

### *3.1. Poly-L-lysine synthesis*

The monodisperse polymers of L-lysine (of 7, 8, and 20 residues) were manually synthesised by using Fmoc chemistry on 0.22 mmoles of Rink amide 4-Methylbenzhydrylamine hydrochloride copoly (styrene-1% Divinylbenzene) resin (100-200 mesh, loading 0.74 mmol/g, Merck-Millipore). Synthesis was carried out only with L amino acids protected with N-fluorenylmetoxycarbonyl group (Fmoc). The resins were solvated by successive washing

steps with dichloromethane (DCM, 3×1 min), 37% piperidine and 0.07% Triton®X-100 in DMF (3×5 min), N,N-Dimethylformamide DMF (6×1 min), and DCM (3×1 min). Fmoc protected amino acids (Merck-Millipore) were incorporated using standard procedures and N,N'-Diisopropylcarbodiimide / 1-Hydroxybenzotriazole hydrate (1:1) in DMF (0.3 M of Fmoc-amino acid) as coupling agents. Fmoc group was removed by treatment with 37% piperidine and 0.07% Triton®X-100 in DMF. After each deprotection, peptidyl resins were washed with DMF (7×1 min) and DCM (2×1 min).

A ninhydrin test was used to monitor amino acid couplings and deprotections along the dipeptide synthesis [11]. Finally, the peptides on Rink-MBHA resin were cleaved by treatment with 3 ml of a mixture of Trifluoroacetic acid/water (95:5 v/v) at 0°C for 15 min and at room temperature for 1.5 hrs. The acid was removed by evaporation and the crude peptide was precipitated with tert-butyl methyl ether/cyclohexane mixture (1:1). The products were characterised by HPLC-MS and ESI mass spectrometry. UV detection was performed at 220 nm, in an Agilent Technologies SQ LC/MSD G6120B instrument. The polymers were purified by HPLC using a Symmetry C18 Preparative Column (100 Å, 5 $\mu$m, 19 mm × 150, Waters®) in a Waters instrument with a 2487 Dual $\lambda$ Absorbance Detector, Waters 600 controller, Waters Delta 600 pump, degasser and Waters Fraction collector III with a manual injector Rheodyne®3725i controlled by the Empower II software.

### *3.2. Experimental techniques*

We obtained a series of UV and CD spectra of our peptides in aqueous solution. The types of spectra were simultaneously acquired using a Jasco J815 spectropolarimeter interfaced with a Peltier unit. Far-UV measurements were performed in 1 mm pathlength quartz cells with PTFE stopper (Hellma). In all cases, three scans were done with a response time of 2 s, a band width of 0.1 nm, and a scan velocity of 20 nm/min.

A range of concentrations between 25-200 $\mu$M was used for poly-L-lysine of 7 and 8 residues solution, 10-100 $\mu$M for poly-L-lysine of 20 residues and 0.25-2 $\mu$M for commercial poly-L-lysine. The experiments in aqueous solution were acquired at 295 K, in NaOH (1 mM, pH 10.8) storaged under nitrogen to avoid pH change during the experiments. For each experiment, a blank solution spectrum was subtracted. The spectra were collected between 187-250 nm. The raw ellipticity was converted to molar ellipticity, $[\theta]$. Samples were left 30 min at 278 K to allow equilibration before each experiment.

## 4. Results and discussion

The Far-UV spectra are a measure of the absorbance of our peptides in solution, which is equal to the extinction coefficient multiplied by the pathlength and the molar concentration. From the analysed spectra of the three synthesised peptides, we found better results with the polylysine of 7 residues. In the inset of Fig. 1, we present the extinction coefficient $\varepsilon$ of this peptide. Assuming an identical polarisability for each residue and using Eq. (11), we obtain the imaginary part of $\alpha$ of a single residue and its real part is computed through Eq. (12). Both components of $\alpha$ are shown in Fig. 1.

In the following stage, the complex polarisability is used together with the structural information of the peptide under study to determine its corresponding ellipticity from Eq. (13). Starting with the optimised polylysines obtained by DFT, we obtain the positions of the centres of polarisability (CP) of each monomer, which are analogous to the centres of mass (instead of using the atomic masses, we employ the dipole polarisabilities of neutral atoms [12]). The electric field $\vec{\mathbf{E}}_i$ is null in the $z$-direction and the vectors $\vec{\mathbf{e}}_i$ come from Eq. (9). In order to improve the statistics of $[\theta]$, we rotate the peptide along a spiral on the unit sphere, as suggested by Park *et al.* [13]. The constants $k_1$ and $k_2$ are determined by comparing our spectra with

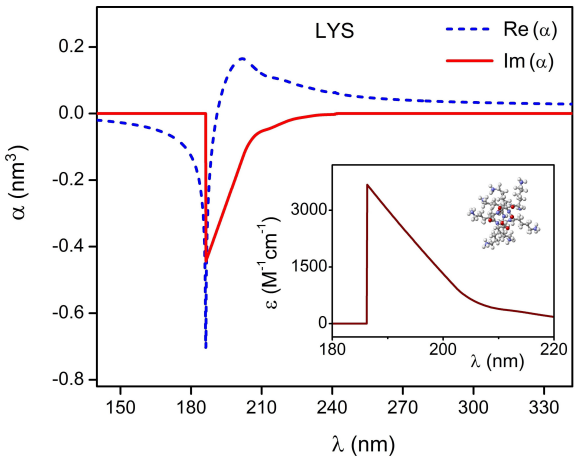



**Figure 1.** Real and imaginary parts of the polarisability $\alpha$ per residue of lysines belonging to a peptide. Both components were extracted from the extinction coefficient $\varepsilon$ of a polylysine of 7 residues and the measured $\varepsilon$ is shown in the inset of the figure (the details of the calculation can be consulted in the text).

the well-established profiles of known structures [7]. As a result, in Fig. 2 we draw the spectra of two polylysines of 150 residues in the configurations of $\alpha$-helix and $\beta$-sheet. As it can be observed in such figure, the two spectra are at least in qualitative agreement with the known spectral shapes.

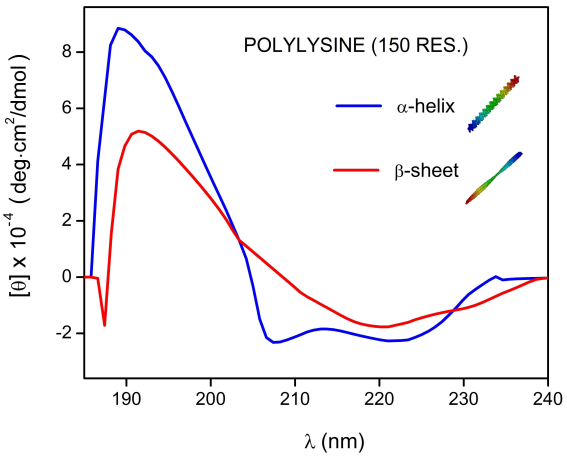



**Figure 2.** Molar ellipticity $[\theta]$ of two polylysines of 150 residues each one that are in different structural configurations, namely, an $\alpha$-helix and a $\beta$-sheet.

We finally performed a comparison with experimental data. The commercial polylysine has an estimate of 230 residues and we carried out the corresponding MD simulation with the package Gromacs with the Gromos 53A6 force-field and using explicit solvent (three-point SPC/E representation of water molecules). In Fig. 3, we compare the experimental results with the prediction of our model for the molar ellipticity, provided that both systems have similar conditions of pH and temperature. As it can be observed in such figure, there is a fairly good agreement between the modelled spectrum and the experimental one, especially in the region of $\lambda > 210$ nm.

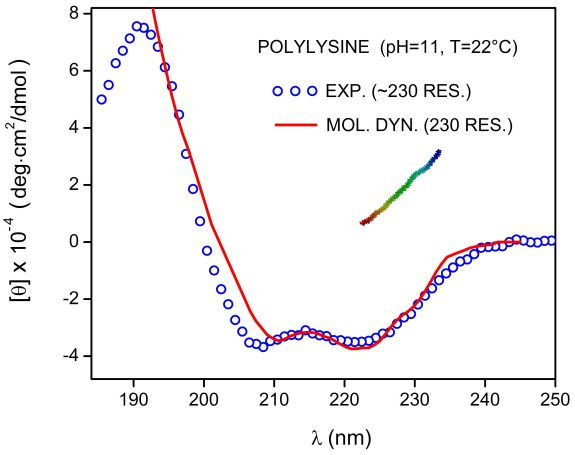



**Figure 3.** Molar ellipticity $[\theta]$ of an experimental system and the resulting spectrum of our model that uses, as an input, a simulated structure of 230 residues.

## 5. Conclusions

We have presented a model of circular dichroism that is able to reproduce, at least in a qualitative way, the known spectra of secondary structures like the $\alpha$-helix and the $\beta$-sheet. In order to calibrate our model, we made use of structures obtained with DFT and experimental data such as the absorptivity and the CD spectra of small peptides that were synthesised by us. These experimental quantities are intended to capture subtle effects that include the interaction between the solvent and the peptide as well as other possible physical interactions (electrostatic, van de Waals, etc.) that still cannot be properly described with QM. More important, the result of our model is in reasonable agreement with an experimental CD spectrum and it was obtained in a reduced computational time.

## Acknowledgments

We appreciate the aid of D. Jacinto-Méndez for providing the MD simulation performed at the cluster Xiuhcoatl of Cinvestav. This work was partially supported by CONACyT through Grant No. 152532. C.G.G.-R. wants to thank the programme *Tiempo de Volver* of Colciencias (Colombia) and C.G.B-C. thanks SIP-IPN Grant No. 20160239 and Medix.

## References

[1] Parr R G and Yang W 1989 *Density-Functional Theory of Atoms and Molecules* (New York: Oxford University Press)
[2] Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864
[3] Sakurai J J 1994 *Modern Quantum Mechanics* (New York: Addison-Wesley Publishing Co.)
[4] Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133
[5] Hoičevar T and Demšar J 2014 *Bioinformatics* **30** 559
[6] Stephens P J, Devlin F J, Chabalowski C F and Frisch M J 1994 *J. Phys. Chem.* **98** 11623
[7] Tinoco I *et al.* 2014 *Physical Chemistry* (Boston: Pearson Education Inc.)
[8] Besley N A and Hirst J D 1999 *J. Am. Chem. Soc.* **121** 9636
[9] DeVoe H 1965 *J. Chem. Phys.* **43** 3199
[10] Miller K J 1990 *J. Am. Chem. Soc.* **112** 8543
[11] Sarin V K *et al.* 1981 *Anal. Biochem.* **117** 147
[12] Anderson K and Sadlej A J 1992 *Phys. Rev. A* **46** 2356
[13] Park S, Bardhan J P, Roux B and Makowski L 2009 *J. Chem. Phys.* **130** 134114