University of South Bohemia in České Budějovice

Faculty of Science

Theoretical study of bacteriochlorophyll aggregates using methods of quantum chemistry and molecular mechanics

Bachelor thesis

Marc Ninou Codina

Supervisor: Mgr. David Řeha, Ph.D.

České Budějovice 2019

Ninou Codina M., 2019: Theoretical study of bacteriochlorophyll aggregates using methods of quantum chemistry and molecular mechanics. Bc. Thesis, in English 40p., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.

Annotation:

This bachelor thesis is focused on the theoretical study of bacteriochlorophyll aggregates. Bacteriochlorophyll (BChl) is a photosynthetic pigment found in light-harvesting antenna systems called chlorosomes. It has been found, that in chlorosomes these aggregates have a lamellar nature, and form distinct curvature. The aim of the thesis was to attempt to identify the origin of curvature formation in chlorosomes aggregates from a local energy perspective, by geometrically optimizing the system and calculating the local interaction energy of tetramer structures located in different sections of BChl aggregates employing different methods of computational chemistry. The results suggest that curvature does not originate from local interactions.

Declaration:

I hereby declare that I have worked on my bachelor's thesis independently and used only the sources listed in the bibliography. I hereby declare that, in accordance with Article 47b of Act No. 111/1998 in the valid wording, I agree with the publication of my bachelor thesis, in full to be kept in the Faculty of Science archive, in electronic form in publicly accessible part of the STAG database operated by the University of South Bohemia in České Budějovice accessible through its web pages. Further, I agree to the electronic publication of the comments of my supervisor and thesis opponents and the record of the proceedings and results of the thesis defense in accordance with aforementioned Act No. 111/1998. I also agree to the comparison of the text of my thesis with the Theses.cz thesis database operated by the National Registry of University Theses and a plagiarism detection system.

České Budějovice, 09.12.2019

……

Marc Ninou Codina

Acknowledgments:

I would like to express my deepest appreciation to my supervisor Mgr. David Řeha, Ph.D. for his outstanding guidance, help and support during this work. I couldn't be luckier to have been introduced to computational chemistry by him. Also, I would like to thank Doc. Martin Kabeláč, Ph.D. for the help given during this work and for always having his door open for me.

I would like to express my gratitude to my family for their support and patience besides my admiration for the adaptation and experience these past three years have been for us, what an amazing adventure it has been so far.

In addition, I would also like to thank Dr. Roman Tuma for allowing me to work on this topic and the whole laboratory community, for making a friendly environment in which working was enjoyable and easy.

List of abbreviations

AFM	Atomic Force Microscopy
BChl	Bacteriochlorophyll
BO	Born-Oppenheimer
BSSE	Basis Set Superposition Error
CC	Coupled Cluster theory
CI	Configuration Interaction
CP	Counter Poise Correction
Cryo-EM	Cryo-Electron Microscopy
DFT	Density Functional Theory
EM	Electron Microscopy
GTO	Gaussian Type Orbitals
HF	Hartre Fock
MM	Molecular Mechanics
MO-LCAO	Molecular Orbital-Linear Combination of Atomic Orbitals
MPn	Møller Plesset Perturbation theory
NMR	Nuclear Magnetic Resonance
QM	Quantum Mechanics
QM-MM	Quantum Mechanics-Molecular Mechanics
SCF	Self-Consistent Field
SP	Single Point
STO	Slater Type Orbitals
TEM	Transmission Electron Microscopy

Table of contents

1. Introduction

1.1 Bacteriochlorophyll

Bacteriochlorophyll (BChl) is a photosynthetic pigment found in light-harvesting antenna systems called chlorosomes. Bacteriochlorophyll can be found in five forms BChl a, b, c, d and e. BChl a and b are mainly found in the reaction center of chlorosomes with other proteopigments but these do not assemble to form aggregates. On the other hand, BChl c, d, and e due to their unique structure, are capable of self-assembling into higher-order structures densely packed, having tight excitonic coupling and enabling their host organism to live and perform photosynthesis in poor light environments [1]. It has been proposed that these pigments could be used in nanotechnology or optical applications including solar energy utilization.

1.2 Chlorosomes

Chlorosomes (Fig. 1) are extramembranous organelles which work as peripheral antenna systems having the ability to capture light energy extremely efficiently and transport it where needed, allowing host organism to live in poor light environments. Up to date, they are the biggest supramolecular antenna observed. First found in sections of cells from *Chlorobi* species by Cohen-bazire et al [2]. They have been found in three phyla of green photosynthetic bacteria, such as filamentous anoxygenic phototrophs of the phylum *Chloroflexi* (earlier green non-sulphur bacteria), being their main light-harvesting apparatus. They have also been found in the green Sulphur bacteria phylum *Chlorobi* [3]. *Chlorobaculum tepidum* is a photosynthetic bacteria, which has been one of the most studied chlorosome containing species. It has also been the first chlorosome containing bacterium to have its full genome sequenced [4]. Due to having similar chlorosomes but differences in the rest of their photosynthetic apparatus, horizontal gene transfer has been thought to be responsible for the presence of chlorosomes in the three unrelated phyla [5].

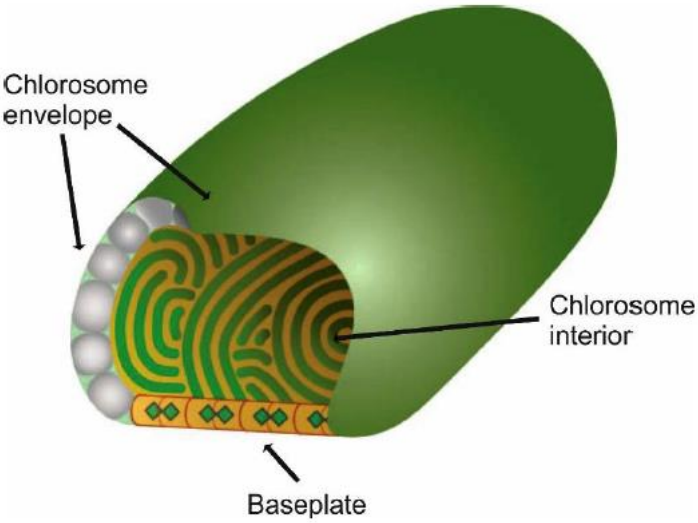

Figure 1. Schematic showing shape and basic structure of a chlorosome (Psencík et al.) [6].

At first, it has been speculated that the envelope of the chlorosome was a lipid monolayer (Fig. 1). Nowadays it is thought that the envelope of the chlorosome is primarily formed by proteins with a small number of lipid molecules in between the gaps.

The chlorosome interior is composed of many BChl molecules (>100.000), that are bound from one side by a proteinaceous baseplate which also contains BChl a and carotenoids and is in charge of transferring the excitation energy onwards to the reaction center. Within the core, aggregates of molecules of BChl c, d and e form lamellar layers [6].

1.3 Structure of Chlorosomes

The shape and size of chlorosomes has been studied extensively using transmission electron microscopy (TEM). It has been shown that both shape and size vary widely between species and exhibiting dependence on growth conditions (Fig. 2). *Chlorobaculum tepidum* has been reported to have chlorosomes with ellipsoidal form, with dimensions around 150 - 200 x 50 x 25 nm. Chlorosomes of other species can differ in dimensions up to a factor of five. Irregularly shaped chlorosomes with rough surfaces were also reported in different species using atomic force microscopy (AFM) [7].

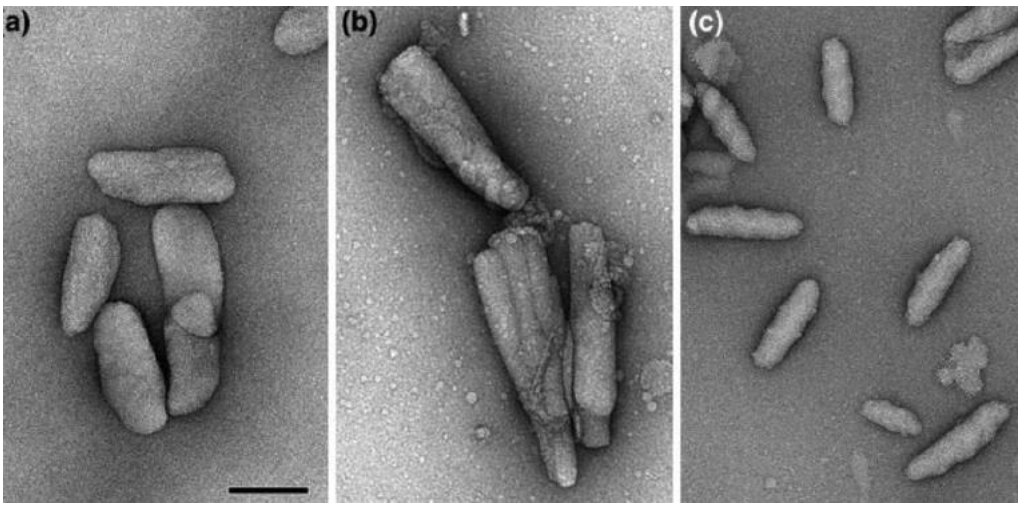

Figure 2: Cryo-Electron microscopy imagines of chlorosomes. The scale bar is 100 nm (Oostergetel et al.) [8].

(a) ellipsoidal form chlorosomes from *Chlorobaculum tepidum*, (b) conical chlorosomes from *Chlorobaculum tepidum* triple mutant and (c) irregular shape chlorosome from *Candidatus Chloracidobacterium thermophilum*.

Within the chlorosome core 200.000 to 250.000 BChl c/d/e molecules can be found [9]. Chlorosomes, dry weight-wise, are about 50% BChl c, d or e, 30% protein and 10% lipids. The remaining are carotenoids, BChl a and quinones. Thus, the pigment content of chlorosomes exceeds other photosynthetic complexes [10]. The work by Staehelin et al [11,12] pointed that chlorosome core can be bound to a lipid-protein layer approximately 3nm thick; this claim was later confirmed by Cryo-Electron microscopy [8]. It has been suggested that chlorosomes might have a lipid monolayer rather than a lipid bilayer, due to the lack of typical structure within all the different organisms. Despite this, estimates have pointed that the lipid content is sufficient to cover up to 5% of the surface. The rest consist of proteins with smaller polar lipid molecules which fill the void spaces in between [13], with their polar heads pointed towards the cytoplasm [14].

Chlorosomes as shown in Fig. 1, contain a baseplate, which plays an essential role, transferring the excitonic energy from the chlorosome to the bacterial reaction centers found in the plasma membrane [15]. It has been suggested that the large structures found in BChl aggregates within the chlorosome core could be due to the well-ordered baseplate [16,17]. The baseplate is mainly composed of BChl coordinated by CsmA protein and forming up to 50% of the protein mass in chlorosomes. It has been shown, that CsmA is the only protein required for chlorosome formation [5].

Photosynthetic BChl (c, d, e) pigments, are found inside the chlorosome interior, being a crucial factor for the function of the chlorosome. They are capable of self-organizing into higher-order structures, maintaining distances and orientations between them without the presence of protein complexes. Also, carotenoids, quinones and in some thermophilic species, non-polar lipids can be found in the chlorosome interior [18,19].

1.4 Molecular Organization of BChl molecules

BChl c, d and e are only found in chlorosomes. In spite of their name, they are not based on Bacteriochlorin, like BChl a, which is found in the baseplate. The reason for this is due to the fact, that they were isolated from bacteria and named before there was knowledge of their chemical structure [19]. BChl c, d and e are based on the chlorin ring instead (Fig. 3) [19].

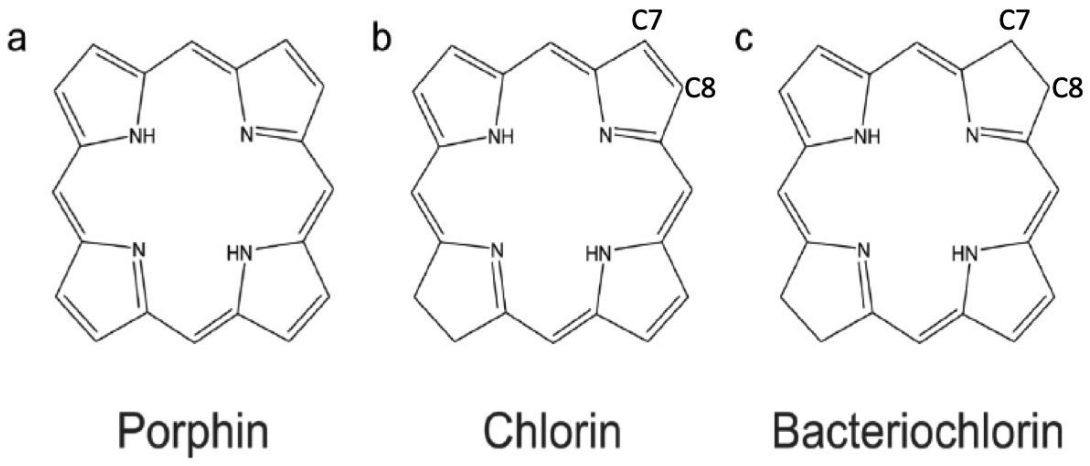

Figure 3: Chemical structure for porphin, chlorin and bacteriochlorin (Daiana K. Deda, et al.) [20].

Chlorin rings are typically found in plant chlorophylls, the difference between bacteriochlorins and chlorins is in the presence of the double bond between C7 and C8 in the latter (Fig. 3). The difference between BChl c, d and e, as shown in Fig. 4, is located at atom C20, where BChl c and e have an additional methyl group and C7 where BChl e has an additional aldehyde group (instead of just the methyl group). There's a fourth possibility, BChl f, which has a hydrogen at C20 and an aldehyde group at C7 [21,22].

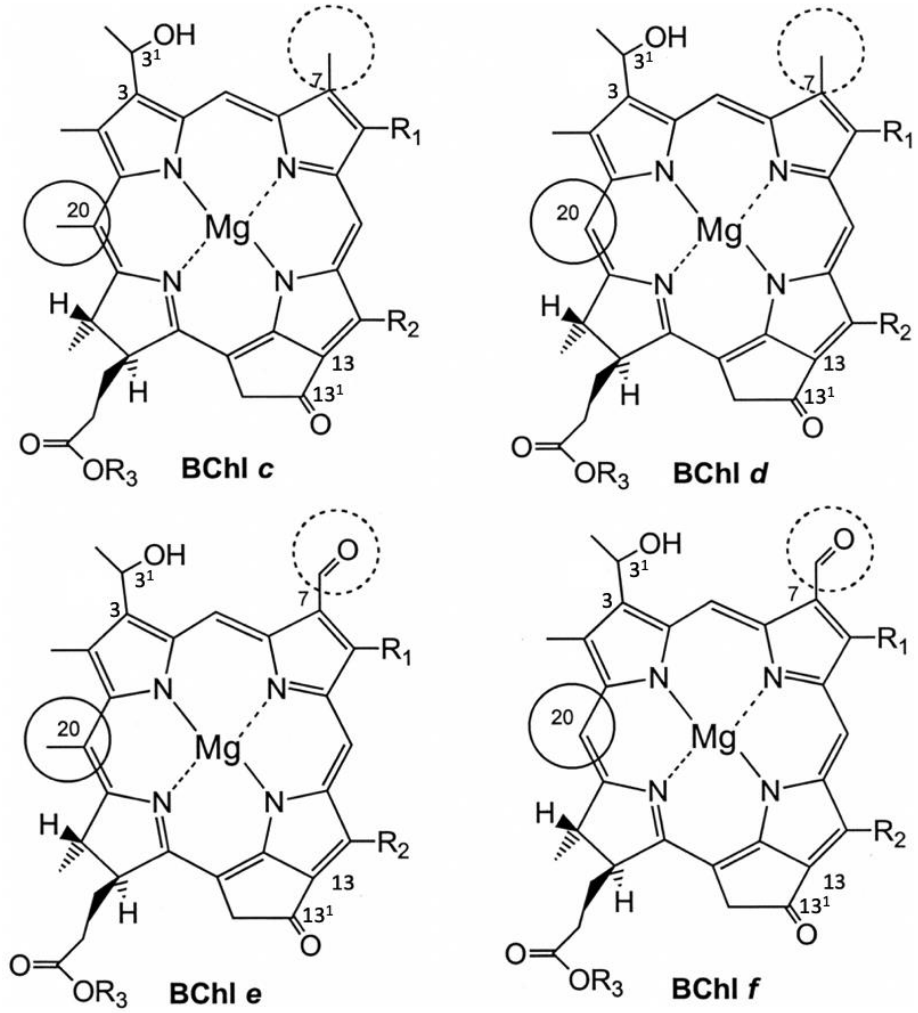

Figure 4: Structure of Bacteriochlorophyll c, d, e and f. Circles denote differences between them. R_1 could be ethyl, n-propyl, iso-butyl or neo-pentyl; R_2 could be methyl or ethyl; and R_3 is in most of the cases, farnesyl. (Robert E.) [23].

BChl c, d and e are able to self-assemble into aggregates due to the hydroxyl group found at asymmetric $C3^1$ atom and the absence of a bulky carboxymethyl group at C13 (Fig. 4) [23]. The presence of short-range interaction among the different chlorin rings assists the self-assemble of BChl c, d or e as a result of their unique chemical structure [10].

These aggregates are coordinated by a Mg^{2+} cation at the center of their chlorin ring, forming coordinate bond with a hydroxy group found at $C3^1$ of another neighbouring BChl (Fig. 4, 5). It is thought that this molecular interaction is key to make possible aggregation. The hydroxy group at $C3^1$ additionally forms the hydrogen bond with the keto group found at $C13^1$ of a third BChl molecule (Fig. 4, 5). Even though this arrangement was accepted, currently there has been a debate over the presence of the H bond, proposing instead a weak coordination between the Mg ion of another BChl to the keto group found at $C13^1$ of the secondary BChl [24].

In spite of the existence of short-range interactions, there are many feasible configurations for BChl molecules within the building block of the aggregate that fulfil NMR derived constraints and short-range interactions [25-32]. From all the possible models, three of them at first got most of the attention (parallel monomer model [28,30], anti-parallel piggy-back dimer [25,26,32] and syn-anti parallel dimer model [27]). The monomer like model is unlikely as there's a requirement for opposite directions for the hydrophobic esterifying alcohol chains that are responsible for the stabilization of the system [33]. However, this last condition is met by both models (syn-anti parallel dimer and anti-parallel piggy-back dimer).

For the different structural motifs, two types can be built: sheet structures and chain structures. The main difference is, that in the sheet structures the bonds between Mg^{2+} atoms and $C3^1$ hydroxy group are formed for all of the stacked BChl-c layers (Fig. 5). On the other hand, in the chain structures the bonds between Mg^{2+} atoms and $C3^1$ hydroxy group are formed only between the two BChl layers forming the BChl layer pairs. These structural models can be divided by the mutual orientation of dipole moments within the dimers into parallel and antiparallel motifs. Combining these, a system of four groups is obtained for the different combinations: (Parallel chains, Antiparallel chains, Parallel sheets and Antiparallel sheets), shown in Fig. 5. [34]

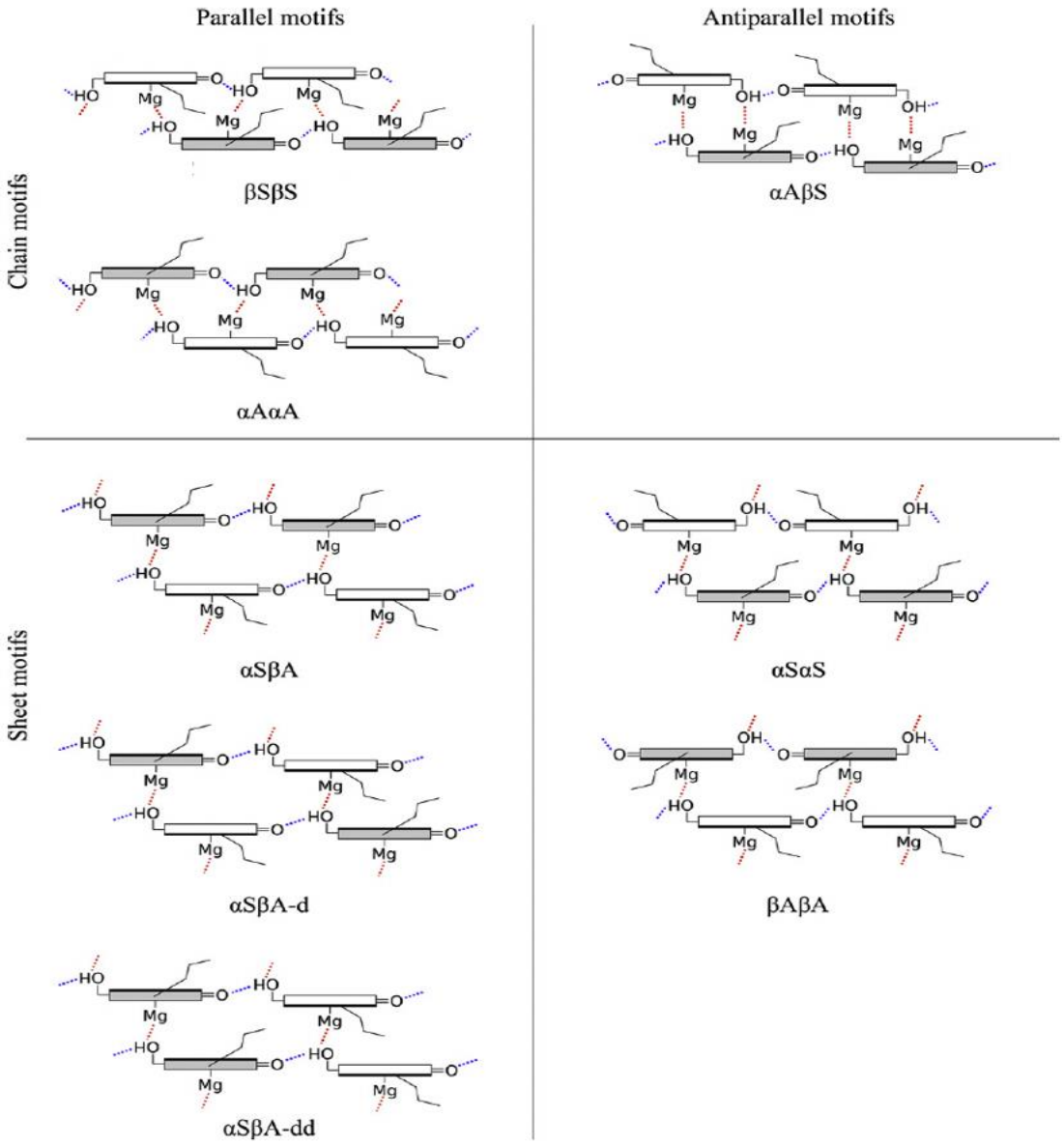

Figure 5: Structural motifs of BChl c divided in (chain and sheet) and by mutual orientation of dimers (parallel and antiparallel) investigated by Alster et al. [34]. A shaded rectangle means that the molecule is in a configuration where C17 is protruding forwards while a white empty rectangle means the molecule is in the opposite orientation.

From the possible eight configurations of BChl-c molecules, which take into account the different experimental findings and the bonding arrangements found previously, the highest stabilization energy has been found (by Alster et al. [34]) for βAβA and αSβA-dd motifs. Therefore, we have chosen those two plausible structural motifs for our further study.

The organization of BChl aggregates into higher-order structures has been subject of much debate. Early free-fracture electron microscopy (EM) work presented striation patterns and rod-shaped structures in *Chloroflexus aurantiacus* and *Chlorobium limicola* [11,12]. Similar striation pattern features were observed by Cohen-bazire et al. [2], who also was the first person to observe chlorosomes.

Studies with electron micrographs and X-ray scattering observed 2nm spacing for chlorosomes, with 5-10 nm rods, which was interpreted by Psencik et al [35] as the distance between the lamellae oriented parallel to the long axis of the chlorosome, because the length of striations prevailed over the extent of chlorosomes (Fig. 6). Further Cryo-EM tilt series experiments proved, that the lamellar system is not planar but curved. This led to the proposal of undulated lamellar arrangements of pigments for three *Chlorobi* species [35].

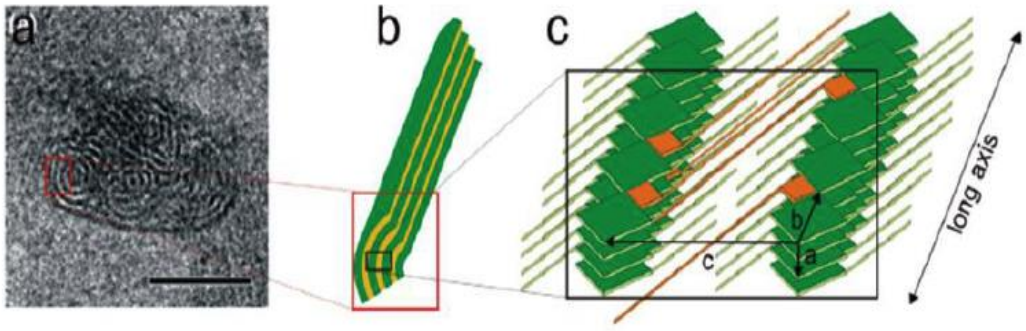

Figure 6: Curved lamellar system in chlorosomes.

(a) View of chlorosome form from EM, (b) schematic interpretation over form pictured, (c) schematic of BChl stacks displayed in green and orange, BChl c and carotenoids respectively. (Psencík et al.). [6]

Studies with cryo-EM performed on *Chlorobaculum tepidum* wild type and bchQRU mutant provided an important insight of lamellar arrangement of BChl aggregates. Upright images of intact chlorosomes, fixed in a thicker layer of ice, showed the distribution of BChl layers and their packing. This illustrates the existence of disordered multilamellar tubular structures which have their diameter between 10 to 30 nm, with non-tubular curved lamellae in between [36].

These multi-lamellar tubular domains extended over the most of the chlorosome, being implemented within the chlorosome in a less ordered smaller curved lamellar matrix. This confirmed and was in agreement with previous studies performed by freeze-fracture experiments and X-ray studies.

As mentioned previously, theoretical model structures have been assembled using the constrains based on experimental observations and insights in bonding patterns which resulted in eight possible structural motifs, shown in Fig. 5, proposed by Alster et al. [34]. These theoretical models have been used for the computational study of BChls layers using molecular dynamics (MD) simulations in order to explain the curvature formation observed experimentally by a PhD student from the research group (Fig. 7).

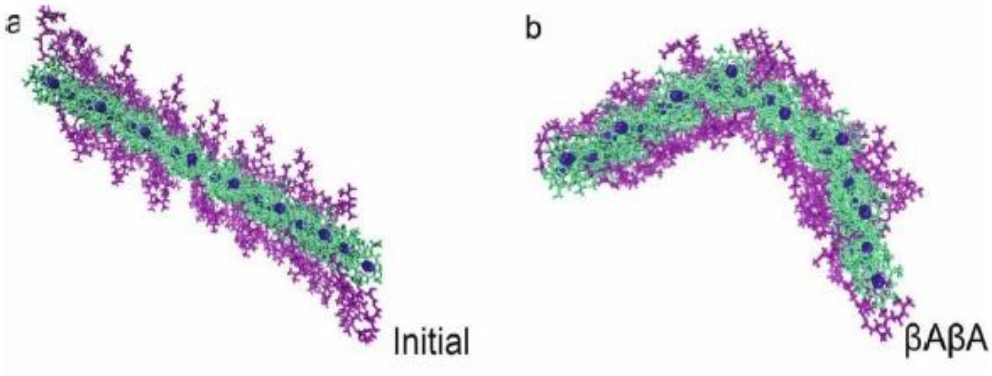

Figure 7: Initial and final configuration for molecular dynamics simulation on sheet structures of BChl aggregates based on βAβA arrangement. (a) Initial starting structure (b) structure after 0.02ns of simulation [37].

As shown in Fig. 7, just 0.02ns of the MD simulation leads to the curved BChl layer, similar to the ones observed by EM. Both experimental and theoretical evidence suggests, that curvature is a natural and spontaneous arrangement for BChl layers. However, the origin of curvature formation still remains unknown. The computational study of local interactions between BChl molecules in the aggregates by Alster et al. [34] does not consider global curved geometry and only the flat local configurations were examined. The geometry for the quantum mechanical (QM) calculations were derived by local energy minimization using molecular mechanics without exploring wider conformational space. Thus, only the local minima (still corresponding to the flat arrangement) were used for the calculations.

The idea behind this thesis is to utilize the curved models of BChl aggregates as results of MD simulations [37] and use them as starting structures for the investigation of local interactions between BChl using QM calculations similarly as in the work of Alster et al. [34].

2. Aims of the study

The aim of this thesis was to investigate the origin of curvature within BChl c pigment layers, which we believe could be encoded in the local interactions between BChl molecules. This was done by utilizing the models of curved BChl aggregates taken from MD simulations [37] and comparing the local interactions in the flat and curved parts of the curved BChl aggregates. The computational level used for the local interaction energy calculation by Alster et al. [34] has been improved by introducing hybrid quantum mechanics/molecular mechanics (QM/MM) methods. The main aim of this thesis was to calculate and compare the local interaction energy between BChl molecules located in flat and curved sections within BChl aggregates in order to determine if the system curvature originates from local interactions and can be justified from a local energy perspective. The specific aims of this work were:

1. To locally optimize geometry of bacteriochlorophyll tetramer in the flat and curved sections of curved models of bacteriochlorophyll aggregates (derived from MD simulations) using QM/MM methods.

2. To calculate the local interaction energies between bacteriochlorophyll molecules of optimized structures for different geometrical patterns and to analyse the results.

3. Materials and methods

This section of the thesis discusses the background behind computational methods used during this work. Covering molecular and quantum mechanics alongside basis set and geometry optimization. In addition to what has been done during this study.

3.1 Computational Chemistry

Chemistry has been considered throughout history an experimental based subject, but for the last four decades this is not valid anymore. Chemical properties have been possible to determine through theoretical prediction thanks to efficient and precise modeling alongside the rapid increase in computational power [38].

This theoretical branch of chemistry allows the study of chemical properties and structures at molecular level. Most approaches and methods used in computational chemistry are dependent on the calculation of the energy of the studied system. The methods for calculation of the system energy can be separated in two main groups: quantum mechanical and classical molecular mechanics.

Quantum mechanical methods are based on the quantum theory describing the system at the electronic level. Since these methods describe the electronic structure of molecules, they are convenient for the determination of intermolecular interactions, modelling of electronic properties, calculation of spectral properties and changes in covalent structure, e.g. chemical reactions. Due to high computational demands, the method has a limit for the size of the system studied.

Molecular mechanics is based on the simplified mechanical models describing system at the atomistic level. This way, we can study large systems with up to millions of atoms, such as large biological assemblies and biomolecules, like nucleic acids or proteins and utilize the principles of Newton's mechanics in order to study their statistical properties (for example using the molecular dynamics simulations). Molecular mechanics even though being faster than quantum mechanics, is not able to predict or reproduce properties based on electron dependence or chemical reactions, where covalent bonds are created or broken.

3.2 Molecular Mechanics

Classical mechanics is the basis for computational chemistry methods referred as molecular mechanics (MM), which was developed for the study of large molecules, adapting the basic principles of Newton mechanics and applying them to atoms and molecules. Hence, avoiding the requirement to solve Schrödinger's equation (1) The potential energy of the system is approximated as a parameterized function of the position of atoms only, utilizing the knowledge of the chemical structure of studied molecules. The potential energy of the system is the sum of the bonding and non-bonding interactions between atoms. The bonding interactions are based on the mechanical model of the covalent bonds (modeled as "springs") and their calculation require empirical parameters for each covalent bond. The non-bonding interactions are based on atom pair interactions, both electrostatic (modeled by Coulomb potential) and van der Waals interactions (modeled by Lennard-Jones potential). The calculation of non-bonding interaction requires empirical parameters for atomic charges (can be also derived e.g. by QM calculations) and Lenard-Jones parameters. This MM approach is very useful when large macromolecules such as nucleic acids or proteins containing very large numbers of atoms are studied. The MM methods require a fraction of the time for the calculations of the given system in comparison with high-level computationally demanding QM methods, however and, they are not capable to predict or reproduce properties that are electron dependent [39,40]. As MM methods require prior knowledge of the chemical structure of the system, they cannot model formation or breaking of the covalent bonds, thus they cannot be used to study the chemical reactions.

3.3 Molecular Dynamics

Molecular dynamics is a computational technique, which, was originally developed for modeling of liquids [41], but has proved to be a useful tool for the modeling of complex biomolecules (see e.g. Schlick et al.) [42]. Molecular dynamics simulations allow the study of the systems over time. The molecular motion in a system is performed according to Newton's second law - the force equals to rate of change of momentum. When employing this to all the atoms (molecules) of a system, we obtain the trajectory as a collection of coordinates for every atom at a certain time. The force is calculated from the potential energy of the studied system which in turn depends on the position of all particles. In the very large systems this leads to the so-called many-body problems and the equations of motion cannot be solved analytically.

It must be solved numerically using integrators or/and finite difference approximations to solve the equations of motion [43]. In order to take in to account the fastest motions such as bond vibrations, the time step for integrators must be very small (usually 1-2fs), thus MD simulations require an enormous number of calculations. In order to facilitate a large number of calculations, MM methods are usually employed for calculation of potential energy of the system.

3.4 Quantum Mechanics

Quantum mechanics (QM) is considered to be the most precise tool to describe nuclei and electron physical behaviour. Theoretically it is capable to predict any properties for individual atoms or molecules [44], however it requires to solve the Schrödinger equation (1). The analytical solution of Schrödinger equation is limited for one-electron systems only, e.g. He^{+} or Li^{2+}. Approximations are required when the system is larger.

$$\hat{H}\Psi = E\Psi \tag{1}$$

Being Ψ the wave function (eigenfunction), $\hat{H}$ being the Hamiltonian operator and E the total energy of the system. The Hamiltonian operator describes the total energy of the system and it is the sum of the potential energy operator and kinetic energy operator, which is derived from another operator called Laplacian. The Hamiltonian operator is applied to all particles of the system, electrons and nuclei. The Born-Oppenheimer (BO) approximation is an assumption which separates the motion of the nuclei and electrons by recognizing the difference in mass between electrons and atomic nuclei [45]. By application of BO, the simplified Schrödinger equation treats the kinetic energy of electrons, the attraction between nuclei and electrons as well as repulsion between electrons [46]. Due to the electron-electron repulsion term (pair interaction) it is impossible to simplify the Schrödinger equation to one-electron form and thus it is impossible to solve it exactly. Therefore, additional approximations must be introduced. Based on those additional approximations, quantum mechanical methods are divided in the following groups: Semiempirical, *Ab initio* and Density functional theory (DFT). For instance, semiempirical methods use empirical approximations and allow us to study bigger systems up to several hundred atoms, but resulting energies are not accurate enough for all of systems.

However, on the other hand, more accurate *ab initio* and DFT methods e.g. coupled clusters can be effectively used for much smaller systems (up to 100 atoms) due to high computational demand, but they provide more reliable results for wide range of the systems.

The high computational demands of QM methods make it difficult to use them for large organic and bioorganic systems. On the other hand, MM methods are significantly less demanding for CPU cost, and it can be used easily for very large organic and bioorganic systems. However, MM treatment is less accurate, and it fails to accurately describe all interactions (like π- π stacking). MM also fails to predict properties that depend on electronic distribution and cannot be used for modelling chemical reactions (like enzymatic reactions). In order to solve this problem, compromise hybrid QM/MM methods has been developed [47]. The QM/MM methods combine accurate QM methods with less demanding MM methods. QM calculations are applied for a specific region of the system (called QM region), where accurate QM calculations are required (like reaction centre of a protein), while MM calculations are used for the rest of the system (MM region) (Fig 8). Thus, the QM region would be a rather small region of interest of an organic or bioorganic system.

The biggest problem of this hybrid approach (QM/MM) is to adequately model the interface or boundary between the QM and the MM regions. A solution for this is to "cleave" bonds that occur at the interface between the regions and "cap" each end of the bond with a hypothetical hydrogen bond [48].

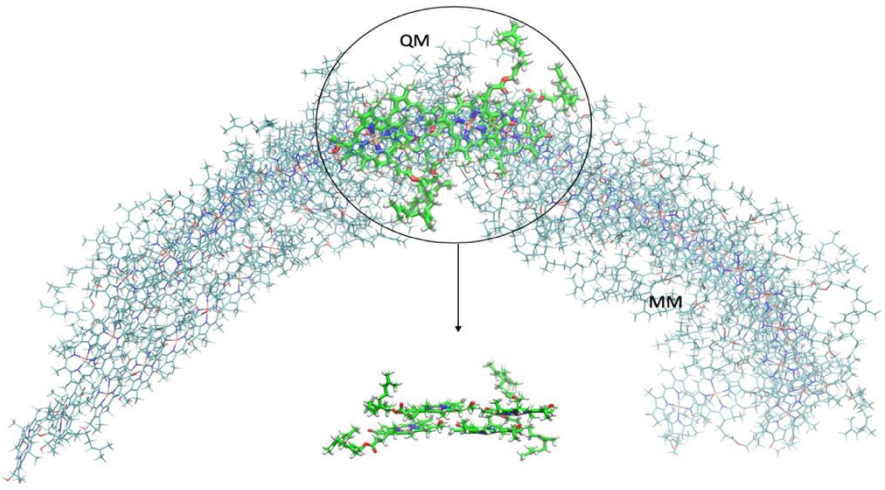

Figure 8: Setup used during this study for energy minimization of BChl tetramer using QM/MM method. QM region depicted inside the circle denoted in green consisting of BChl tetramer including its tails while the rest of the system denoted in blue is the MM region.

3.4.1 Ab initio Methods

Ab Initio from Latin origin means "from the beginning". It is capable to recreate experimental data within certain limits without taking into count empirical parameters unlike molecular mechanics or semiempirical methods. Seeing that, this method is preferred when there's barely or no experimental data available [49].

One of the simplest *ab initio* methods, based on its low-computational cost is Hartre Fock (HF) method. Based on BO approximation, as mentioned above, it separates and treats nuclear and electronic motions of each other individually. The HF methods use a unique approximation (Fock matrix defined by the Fock operator) which is an estimation to a Hamiltonian operator for a quantum system. In this way, the electron-electron repulsion (interaction of electron pair) is reformulated in the one-electron form (one electron interacts with the cloud of the remaining electrons). By this method, the Schrödinger equation is solved for fixed nuclear configurations, obtaining the electronic energy of a system. The downside of this method is the error created due to the lack of electron correlation.

The wavefunction of the HF method is built from monoelectronic functions (orbitals) in the form of the so-called Slater determinant [50] and therefore the correlation energy (dynamical correlation energy) is not included [51]. This energy difference (correlation energy), can be calculated by further methods which are known as post-Hartre-Fock methods. Examples are: Møller-Plesset perturbation theory (MPn), coupled-cluster theory (CC) and configuration interaction (CI). However post-Hartree Fock methods are extremely computationally demanding and can be effectively used only for the calculations of the smaller systems (up to 50 atoms).

Møller-Plesset Perturbation theory (MP) takes into a count electron correlation, as a small perturbation with respect to the result obtained from HF. The HF becomes the first order perturbation, MP2 second, MP3 third, etc. The higher the order of the method the more accurate the result will become even though MP2 is the most used., even though it does slightly overestimate the dispersion contribution. Therefore, it is not suitable in cases of very accurate description of intermolecular interactions [52-54].

3.4.2 Density functional methods

Unlike the methods mentioned above, density functional theory (DFT) works with electron density instead of describing the full wave function. These methods are quite popular and commonly used since they are capable of calculating any molecular property and can be applied to large organic and bioorganic molecules. The formulation by Hohenberg and Kohn [55] also takes into a count a significant part of the correlation energy using approximate DFT functionals. Furthermore, the computational demands of the methods are lower in comparison with other post-Hartre-Fock methods such as MP2. The reason for this, is due to correlation functional, which includes portions of short-range correlation energy, allowing to capture a significant part of the correlation energy with less computational time. However, it does not describe London dispersion energy (long-range correlations), that is normally included in the correlation energy calculated by post-Hartre-Fock methods (like MP2). Therefore, DFT can't describe specific intermolecular interactions (like $\pi - \pi$ stacking) which are usually found in biological systems. This problem can be corrected by adding a dispersion term either to the final energy or already to the DFT functional, e.g. B-97D functional [56]. For these reasons, DFT are one of the most popular computational methods nowadays.

3.5 Basis Sets

In quantum chemical calculations, molecular and atomic orbitals are approximated by expansion into a sum of simple basis functions:

$$\Psi = \sum_i c_i \phi_i \qquad (2)$$

In Eq. (2) functions $\phi_1, \ldots, \phi_n$ are the basis set functions. Molecular Orbital-Linear Combination of Atomic Orbitals (MO-LCAO) is the notation used for this approximation.

Expansion coefficients c_i are numerical values obtained from Self-Consistent Field (SCF) procedures. Basis for atoms contain a minimum of one function for each orbital occupied by electrons. This is called the minimal basis set, for example carbon, nitrogen or oxygen atom requires 5 functions (1s, 2s, $2p_x$, $2p_y$, $2p_z$) in the minimal basis set. However, the minimal basis provides poor description of the system. The greater number of functions in the basis set present, the better description of the system is, for this reason it is good to calculate on double or triple basis set.

The solution of Schrödinger's equation for a hydrogen atom is known, consequently, the hydrogen like wave functions can be used to define ϕ. These functions are then further simplified into the so-called Slater type orbitals (STO) which are used as basis set functions for any other atom. However, it is more common nowadays to use Gaussian type orbitals (GTO) because of lower computational expense during calculations. The downside of GTO is the different behaviour (in comparison to STO) depending on the distance of a region from the nucleus. This can be solved by increasing the number of GTOs, thus one STO is described by the more GTOs. This type of basis set can be found in literature with the notation: 3-21, 6-31, etc. The notation consists of the first number standing for primitive Gaussians (GTO) comprising the STO at the atomic core, the second and third number, indicate the number of GTO for each valance STO, two numbers indicate that valence STO is split in two functions (double basis set).

In order to describe the effect caused by neighbouring nuclei atoms in molecule to electrons of study, additional polarization functions, are included in the basis set. These contain extra functions for the different d and p orbitals. For example, above mentioned atoms (C, N or O) would have 5 additional d-orbital functions ($3d_{xy}$, $3d_{xz}$, $3d_{yz}$, $3d_{x2-y2}$, $3d_{z2}$). This can be found in literature with the notation by * or **: 3-21*, 6-31** [57]. The notation '*' implies added d polarization functions on non-hydrogen atoms only and '**' added d polarization on non-hydrogen atoms and p polarization functions for hydrogen atoms. Another type of basis set was defined by Ahlrichs and co-workers referring to the initial formations of the split valence and triple zeta basis sets from this group (SV, SVP, TZV and TZVP) [58]

3.5.1 Basis Set Superposition Error

QM calculations use finite size basis set. This may cause trouble when one wants to calculate the interaction energy since the method treats the different fragments isolated and describes them by using own suited basis sets for each. However, the calculation of the complex is performed in the larger basis set, as it contains the orbitals for every individual fragment of the system. The interaction energy is then obtained using the equation: $E_{Interaction} = E_{Complex} - \Sigma E_{Fragments.}$ But the interaction energy for the whole system has larger and thus more accurate basis set compared to the individual fragments leading to an overestimation (more negative than it should be) in $E_{Interaction}$. This specific error is called Basis Set Superposition Error (BSSE).

BSSE is, therefore, a mathematical artifact based on the fact that the energy of the complex is calculated with a more complete basis set since it includes all the functions for the different subsystems. Compared to the constituent fragments which are described only by their own basis set. The error decreases when larger basis sets are used, but this may not be feasible depending on the studied system due to slow convergence. This error can be eliminated by using counter poise correction (CP) method [59]. By using this correction, the different isolated fragments are described by the very same basis set as the whole complex. When individual fragment is calculated, the additional basis functions are added to the positions where the atoms other fragments of the complex used to be during the calculation of the complex, thus the same basis set is used for individual fragment as for the complex.

3.6 Geometry

The structure of biological macromolecules is usually determined by different methods such as X-ray diffraction, NMR or lately electron microscopy [60-62]. The theoretical determination of the geometry for a model system is a key, because experimental data may not be available. Unfortunately, X-ray measurements do not provide information over the position for hydrogen atoms and theoretical calculations must be used to calculate the complete geometry. A way to obtain the geometry of the studied system can be geometrical optimization. The concept behind this is to reduce the energy of the system, calculated by a theoretical method (which calculates the energy), while systematically changing the geometry of the system. Steepest descent, conjugate gradient and Newton-Raphson are a few of the different methods one can use for energy minimization. They are based on the knowledge of the gradient (first derivatives) and in the case of Newton-Raphson approach also Hessian (second derivatives), calculated by the theoretical method. The downside of these calculations is the dependence on the original geometry [63].

3.7 Work done during the study

The starting point of this study was the previously done simulations of BChl c aggregates using molecular dynamics by a PhD student from the research group [37] based on structural motifs proposed by Alster et al. [34]. Alster and his co-workers calculated the local interaction energy of BChl molecules within the tetramer on QM level; however, the BChl tetramer was treated as an isolated system and its interactions with surrounding BChl molecules from the aggregates were ignored for the energy calculations.

Surrounding BChl molecules were included only during geometrical optimization (energy minimization) on MM level. The idea for our study was to improve the computational level and include the surrounding BChl molecules from the aggregate to the QM calculation of the local interaction energy of BChl molecules within the tetramer as well as the geometrical optimization of the system. The MM calculations used by Alster [34] does not describe correctly the densely packed systems, where $\pi - \pi$ stacking interactions plays an important role, which is the case of the densely packed arrangement of BChl molecules in studied aggregates. Therefore, it is advantageous to use QM methods that properly describe $\pi - \pi$ stacking interactions. However, due to the size of the system, only QM/MM hybrid method can be employed, where the BChl tetramer itself was treated as the QM region and the rest of the BChl molecules (forming the aggregate) were treated as MM region.

Model structures used in the above mentioned MD study [37] were assembled using 3^1R-[E,M] BChl-c_f (abbreviation explained in Fig. 9), pigment assemblies were constructed for various pigment utilizing dimer as building block units which fulfilled constrains from X-ray scattering and the principal interactions between BChl-c molecules, its coordination of Mg cations by the hydroxyl groups [34]. The built model of a layer of BChl-c aggregate contained 80 BChl-c molecules (10x8 sheet, 10 BChl molecules in stack and 8 BChl molecules bonded by hydrogen bonds).

So that IUPAC rules are followed for tetrapyrroles, S, was adapted to denote "Syn" referring to the arrangement where the hydroxy group at $C3^1$ and C17-propionic acid moiety (Fig. 9) are on the same side of the chlorin ring. Conversely, A, was adapted to denote "Anti" in the case it is at the opposite side. The orientation of the Mg cation with respect to C17 was designed as α and β. Being consistent with the notation proposed by Balaban et al. [64], for the arrangement where the Mg cation is on the same side as C-17 propionic acid moiety β was designated, on the case of being on the opposite face it was classified as α configuration. By this identification, each dimer unit has a four-letter code specifying its stereochemistry. The structures used during this work were βAβA and αSβA-dd, shown in Fig. 10, which as mentioned in the introduction were two structures with highest stabilization energy, therefore, being plausible for the different possible aggregates in chlorosomes [34].

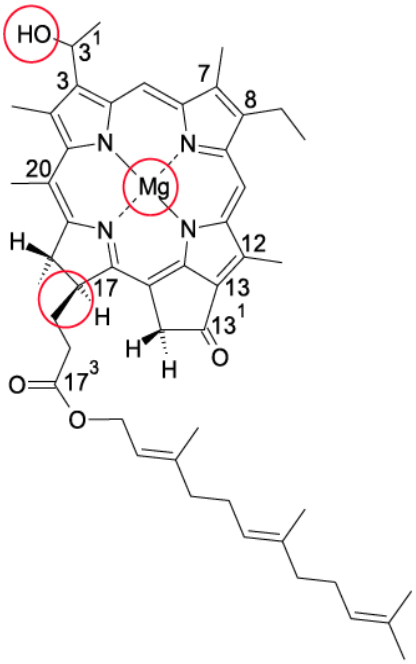

Figure 9. Molecular structure proposed as homolog of BChl-c, 3^1R-[E,M] BChl-c_f. The meaning of the symbols in above abbreviation is following: 3^1R abbreviates for an R configuration in respective chiral center located at $C3^1$, the letters within the brackets exemplify ethyl and methyl as substituents for C8 and C12 respectively. F written in subscript designates the molecule being esterified with farnesyl at C17. The two stereo-centers and the chiral center at C17 (used to define this nomenclature) are denoted with red circles (Alster et al.) [34].

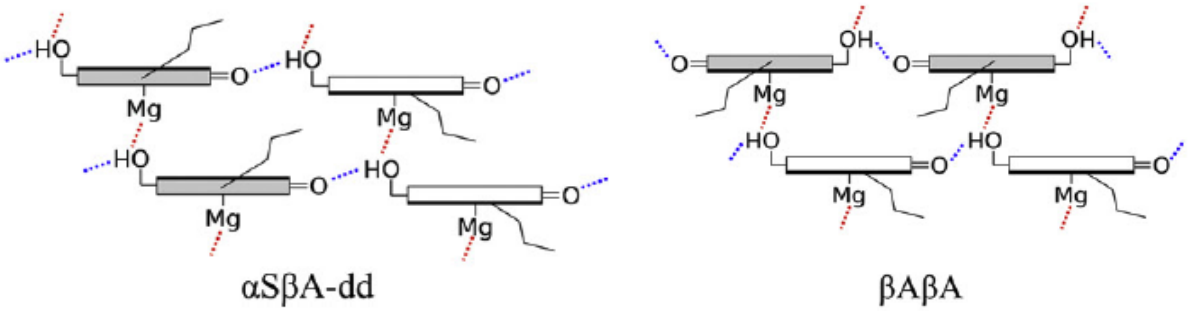

Figure 10. Structural motifs used during this work of BChl-c as they were considered in aggregate models. Shaded rectangle points that the molecule with a configuration with C17 protruding forward while a white rectangle denotes the molecule in the opposite orientation (Alster et al.) [34].

Two BChl tetramers were picked from the curved and flat section of the final curved geometry of the MD simulations for the both studied system (βAβA, αSβA-dd), The BChl tetramer is shown in Fig. 11 and its position within the structure of aggregate is depicted in Fig. 12 and 13.

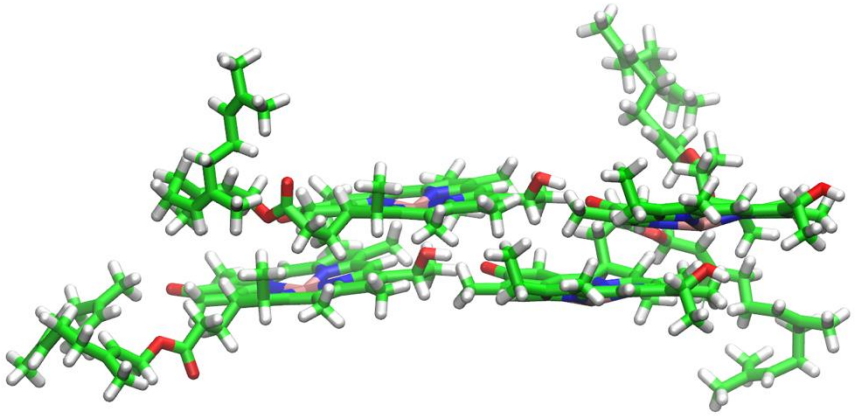

Figure 11. Tetramer structure from αSβA-dd.

The goal of this selection was to compare the local interaction energy between BChl located in both flat and curved local geometries in order to determine if the system curvature can be explained from a local energy perspective, i.e. whether the curvature is an intrinsic property of the BChl c structure and local packing. They were also chosen also in a way to make sure they are surrounded by other BChl molecules of the system to exclude surface effect and mimic the environment.

The geometry of the structures were then individually optimized employing QM/MM hybrid method with the BChl tetramer treated by quantum mechanics and the rest of the system with molecular mechanics, as previously shown in Fig. 8. The calculations were performed using Qsite utility from the Schrödinger software package [65]. For the QM region (BChl tetramer including its tails) the DFT method was used with B3LYP functional and 6-31G* basis set, the MM region (the rest of the system) was described by OPLS2005 forcefield [66].

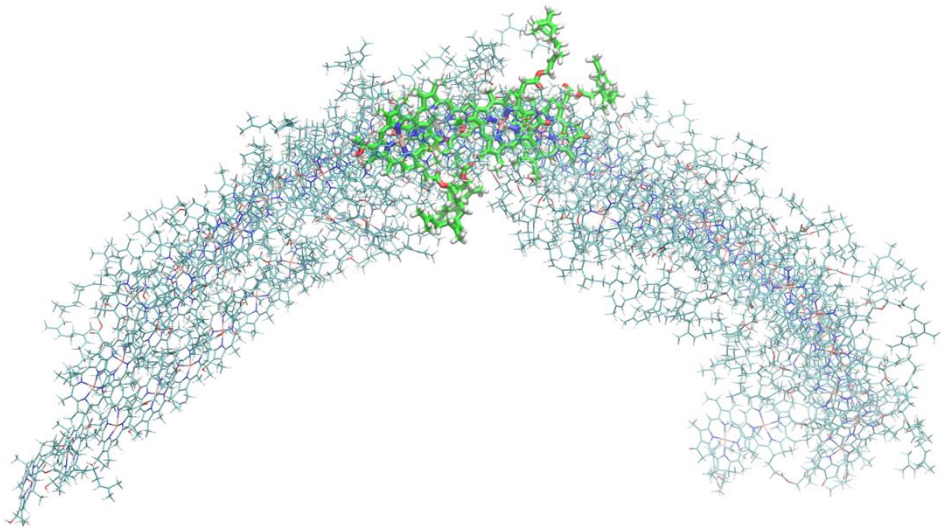

Figure 12: Structural motif αSβA-dd, depicted from above. With chosen tetramer in curved section of the system displayed in green

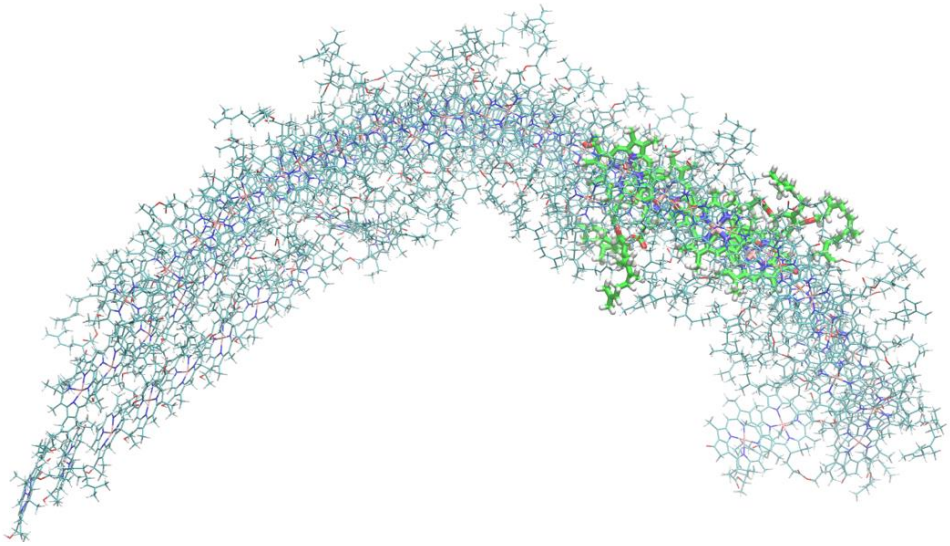

Figure 13: Structural motif αSβA-dd, depicted from above. With chosen tetramer in flat section of the system displayed in green.

The total interaction energy from the geometry optimized tetramers was then calculated using DFT, DFT polarized and MP2 methods done in Gaussian G09 [67]. DFT method with B-97D [56] functional provides reliable results for energy calculations as it also covers correlation energy (treating short-range correlations) and additionally also dispersion term is taken in to account (describing long-range correlations). The interaction energy was derived using a supramolecular approach described by Eq (3).

$$\Delta E_{Tot\_int} = E_T - E_{M1} - E_{M2} - E_{M3} - E_{M4} \quad (3)$$

Where ΔE_{Tot_int} is the total interaction energy of BChl monomers in BChl tetramer, E_T is the tetramer's energy in vacuum, E_{Mx} is the energy of the individual BChl monomer.

In addition, DFT polarized, was used to calculate the contribution of the polarization from the surrounding atoms (used as point charges) to the total interaction energy of the tetramer. The form for this contribution was calculated as follows:

$$\Delta\Delta E_{correction} = \Delta E_T - \Delta E_{M1} - \Delta E_{M2} - \Delta E_{M3} - \Delta E_{M4} \quad (4)$$

$\Delta\Delta E_{correction}$ being the total contribution of the surrounding atoms as point charges to the total interaction energy of BChl monomers comprising BChl tetramer. $\Delta\Delta E_{correction}$ should be added to ΔE_{Tot_int} from Eq. (3). $\Delta\Delta E_{correction}$ is conceptually comparable to dissolvation energy, as it is the energy associated with the "solvation" (formation) of the tetramer within the surrounding, expressed as ΔE_T, subtracting each of its different parts (monomers), in the form of ΔE_{MX}. Therefore, it can be considered as a correction for "solvent", which applied to our studied system would be the atoms surrounding the tetramer.

$$\Delta E_T = E_{(T+chg)} - E_{chg} - E_T \quad (5)$$

ΔE_T being the interaction energy between the tetramer and its surrounding point charges, representing the atoms of the remaining BChl molecules of the aggregate. $E_{(T+chg)}$ being the total energy (tetramer and surrounding point charges), E_{chg} being the energy of the surrounding point charges and E_T being the tetramer's energy in vacuum.

$$\Delta E_{MX} = E_{(MX+chg)} - E_{(chg)} - E_{(MX)} \quad (6)$$

ΔE_{MX} being the interaction energy between the given monomer and surrounding point charges representing remaining BChl molecules (excluding remaining 3 BChl monomers from tetramer), $E_{(MX+chg)}$ being the total energy (monomer and surrounding point charges) and $E_{(chg)}$ being the energy of the surrounding point charges and $E_{(MX)}$ the monomer's energy in vacuum.

Lastly, MP2 method (Møller-Plesset Perturbation Theory) treats electron correlation as a small perturbation with respect to the HF result. MP2 is one of the most commonly used post-Hartree-Fock method, being one of the least computationally demanding (compared to other post-Hartree-Fock methods) options for reliable and accurate estimates of correlation energy. However, due to computational demands, just one calculation was performed in order to confirm and compare the result obtained from DFT.

In addition to the total interaction energy, the interaction energy between the BChl tetramer and the surrounding BChl molecules (forming the BChl aggregate), described in Eq. (7), was calculated by adding the interaction energy of the tetramer with the surrounding point charges representing the rest of the system (Eq. 5) and van der Waals interactions between the tetramer and the rest of system (Eq. 8).

$$E_{Interaction} = \Delta E_T + E_{vdw} \quad (7)$$

$E_{Interaction}$ being the interaction energy between the tetramer and the surrounding BChl molecules, ΔE_T being the interaction energy of the tetramer with the surrounding point charges (Eq. 5) and E_{vdw} being the van der Waals interactions between the tetramer and the rest of the system (Eq. 8).

$$E_{vdw} = E_{Total} - E_{Tetramer} - E_{Rest} \quad (8)$$

E_{vdw} being the van der Waals interaction energy between the tetramer and the rest of the system, E_{Total} being the van der Waals energy of the whole system, $E_{Tetramer}$ being the van der Waals energy of the tetramer and E_{Rest} the van der Waals energy for the rest of the system (BChl molecules in aggregate except for the tetramer).

4. Results

4.1 Optimization

The tetramer structures from the flat and curved section respectively, and for both plausible structures (αSβA-dd and βAβA) were optimized using a hybrid method (QM/MM). The four structures were chosen to be surrounded by the other BChl molecules of the system in order to exclude possible surface effects and mimic the environment. The hybrid QM/MM method allows combination of a high-level theory (quantum mechanics for a specific region (tetramer)) with fast method based on molecular mechanics (MM) for the rest of the system.

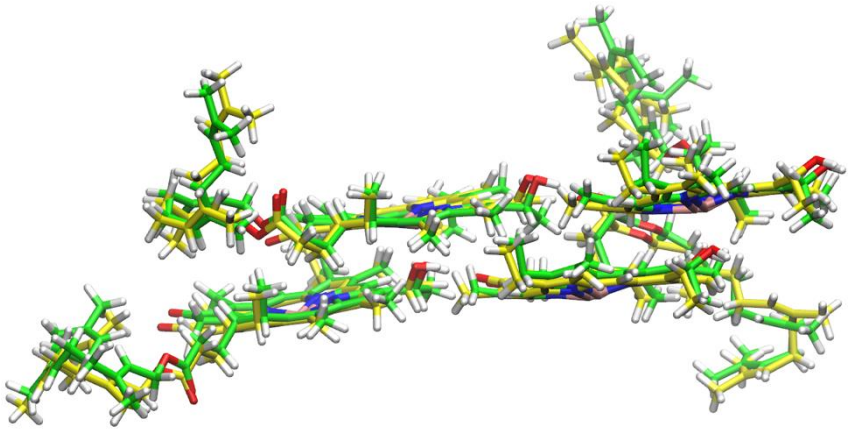

Figure 14: Original and optimized tetramer overlapped. Specifically, tetramer located in flat section of αSβA-dd model. Displayed in green is the original geometry from MD simulation and in yellow is the resulting optimized tetramer.

This means that the tetramer is treated quantum mechanically including the treatment of all electrons, while the rest of the system (remaining 76 BChl molecules) are treated purely by molecular mechanic using all atoms forcefield.

This would reduce the computational demands and allows the system to be optimized within reasonable time. On the other hand, the use of QM method for the specific region (tetramer) allows accurate description of the stacking interactions between BChl molecules of tetramer, that would otherwise would not be possible by using just MM method. The minimization was performed in order to obtain the structure corresponding to the energy minima for the chosen computational method. It should be stressed that the geometrical minima obtained by the method corresponds just to the local minima, since the finding of the global minimum would require larger computational time. The shape of the curvature of the BChl aggregate did not change during the minimization, confirming the fact that just local minimum was found. Further calculations of interaction energy were done on these structures (local minima).

4.2 Interaction energy

The optimized tetramers were then used to calculate the interaction energy employing three different methods: DFT, DFT polarized and MP2. This was done to obtain the total interaction energy of BChl monomers within the tetramer, corresponding to the energy of assembling the tetramer structure from isolated monomers. This energy was calculated for both isolated BChl tetramer as well as the BChl tetramer within the system of all other BChl of assembled layers.

Interaction energy can be theoretically decomposed into different terms, something which is hard or impossible experimentally. According to the Symmetry Adapted Perturbation Theory, interaction energy can be decomposed into four basic components: electrostatic, induction, dispersion and repulsion. The sum of the permanent multipoles interactions between two molecules is considered to be the electrostatic energy. Induction energy assumes the presence of polar (charged) molecules that can create an induced multipole moment in the interacting molecule, resulting in attractive interaction. The dispersion interaction appears from the fluctuation of electron density of a molecule that results in a temporary multipole, generating an induced multipole in the interacting molecule. Repulsion is a destabilizing interaction arising from Pauli exclusion principle, prohibiting two electrons to be in the same quantum state simultaneously [68,69]. However, in our study we have used supramolecular approach in order to calculate interaction energy (Eq 3). Within this approach (as oppose to Symmetry Adapted Perturbation Theory) the interaction energy cannot be decomposed to the individual components.

4.2.1 DFT results

DFT was used due to its accuracy and low computational demands when compared to other methods such as MP2, besides, also taking into account most of the correlation energy. A dispersion term, B-97D [56], was added in order to include long-range correlations (London dispersion energy), which play an important role in our studied systems. TZVP was the chosen basis set, giving the good compromise between accuracy and computational cost. The results obtained are shown in Table 1 for both suitable structural motifs of each optimized tetramer located in the flat and curved section respectively.

Table 1: Interaction energy obtained using DFT for both structural motifs of tetramer structure in flat and curved section

Location Tetramer	βAβA	αSβA-dd	
Flat section	-125,05	-117,49	kcal/mol
Curved section	-108,36	-111,08	kcal/mol

A clear stabilization was observed for the tetramers located in the flat section (i.e. interaction energies are more negative) with respect to the curved sections for both plausible structure motifs. This means that the tetramers found in the curved section are less stable than the ones found in the flat section from a local energy perspective.

4.2.2 DFT Polarized results

To see the effect of the surrounding BChl molecules to the total interaction energy of the tetramer, DFT polarized method was used, in order to include the polarization effect of the surrounding BChl molecules. The calculation of this additional correction due to the point charges of surrounding BChl molecules to the total interaction energies is described in methods and materials, using Eq. (4).

Table 2: Contribution from the point charges of surrounding BChl molecules to the total Interaction energy obtained using DFT polarized method for both structural motifs for tetramer structure in flat and curved section.

Location Tetramer	βAβA	αSβA-dd	
Flat section	-10,19	-9,88	kcal/mol
Curved section	-9,27	-8,13	kcal/mol

The polarizing contribution of the surrounding BChl molecules is more stabilizing (more negative energy) for both structural motifs for the tetramer found in the flat section, following the tendency of the results calculated using DFT. These results were then added to the previous outcomes of DFT in order to obtain the total interaction energies containing the DFT polarization term (Table 3).

Table 3: Contribution from the surroundings added to interaction energy obtained using DFT for both structural motifs for tetramer structure in flat and curved section taking into a count surrounding contribution

Location Tetramer	βAβA	αSβA-dd	
Flat section	-135,24	-127,37	kcal/mol
Curved section	-117,64	-119,21	kcal/mol

4.2.3 MP2 results

MP2 (Møller-Plesset Perturbation Theory) takes into consideration electron correlation, as a small perturbation with respect to the result obtained from Hartre Fock (HF). MP2 is considered as one of the high-level correlated *ab initio* quantum chemical methods even though within the group is one of the least computational demanding methods. Being computationally very effective, estimates reasonably well the correlation energy at a relative low cost and the energies obtained are size consistent. The result obtained is shown in Table 4.

Table 4: Interaction energy obtained using MP2 for the structural motif αSβA-dd for tetramer structure in flat section

Location Tetramer	αSβA-dd	
Flat section	-101,80	kcal/mol

However, the computational demands compared to other methods such as DFT are still high, for this reason, just one calculation was performed in order to confirm and compare the result obtained from DFT. The MP2 calculation corroborated the result obtained using DFT, however, we can conclude that DFT methods overestimates the interaction energy by about 10 kcal/mol.

4.2.4 Interaction energy between tetramer and surrounding

The interaction energy between the tetramer and the surrounding was calculated, as described in materials and methods, using Eq. (7). The results obtained are shown in Table 5.

Table 5: Interaction energy between the tetramer and the rest of the system

Location of the tetramer	βAβA	αSβA-dd	
Flat section	-635,60	-561,56	kcal/mol
Curved section	-618,90	-502,33	kcal/mol

Results show that the interaction between the tetramer and the surrounding BChl molecules of the aggregate is higher (more negative interaction energy) in the flat section than in the curved for both motifs. This points to a similar tendency as observed in the case of the local interaction energy of BChl monomers within the BChl tetramer. However, a much larger difference in the interaction energy between the tetramer and the rest of the system was observed for both studied motifs, with βAβA having significantly stronger interaction between the tetramer and the rest of the system than αSβA-dd.

5. Discussion

By employing electron cryomicroscopy and small angle X-ray scattering it has been found, that BChl aggregates have a lamellar nature in chlorosomes [6, 8, 35, 36, 70]. It has also been demonstrated using electron microscopy that curvature is an essential feature of lamellar layers with arrangements ranging from concentric cylinders in chlorosomes from triple mutant *Cba. Tepidum* to disordered curved lamellae in wild type chlorosomes [36]. Besides, some evidence has been found for the existence of both flat and curved regions in wild type chlorosomes [5].

Present work shows a difference in interaction energy for the tetramer structures (located at flat and curved section within the system respectively) in both plausible structural motifs studied (βAβA and αSβA-dd). However, our results indicate, that from a local energy perspective the BChl tetramer located in the flat section is more stable than the BChl tetramer located in the curved section, for both studied structural motifs, which seemingly contradicts the formation of the curvature within chlorosome layers, observed experimentally and by MD simulations [37]. Thus, the curved nature of chlorosomes cannot be explained from a perspective of the local interaction energy within the tetramer.

When our results obtained using DFT are compared with the calculations performed by Alster et al [34], where the flat layers of BChl aggregates based on the different structural motifs (including βAβA and αSβA-dd) were optimized using pure MM program and single-point (SP) energy calculations of tetramer were performed by QM program for the isolated extracted central tetramer unit, a small difference in the values of interaction energies is found. The values of the interaction energy they obtained were -129,2 and -131,7 kcal/mol for αSβA-dd and βAβA respectively. The cause of this difference can be due to the different geometry of the aggregates (we have used curved BChl aggregates from MD simulation) as well as the fact that the tetramer interaction energy calculated by Alster et al. [34] does not take into account its surroundings.

The MP2 result for αSβA-dd in the flat section corroborated the result obtained using DFT but showed an overestimation of the interaction energy for about 10 kcal/mol by DFT method. Due to high computational demand of MP2 method, just one calculation was performed in order to confirm the result obtained previously.

The interaction energy between the tetramer and the surrounding BChl molecules showed a lower interaction between the tetramer and the surrounding BChl molecules in the curved section than the flat section. Therefore, also seemingly contradicting the natural curvature found in the chlorosomes.

In order to examine and compare the effect of curvature on local geometry, the bacteriochlorin ring of the one monomer (indicated as (1) in Fig. 15) from the tetramer located in flat section was superimposed (RMSD fit) over the same monomer from the tetramer located in the curved section. The resulting overlapped tetramers are shown in Fig. 15, where the monomer number is depicted close to the corresponding Mg cation (displayed in beige).

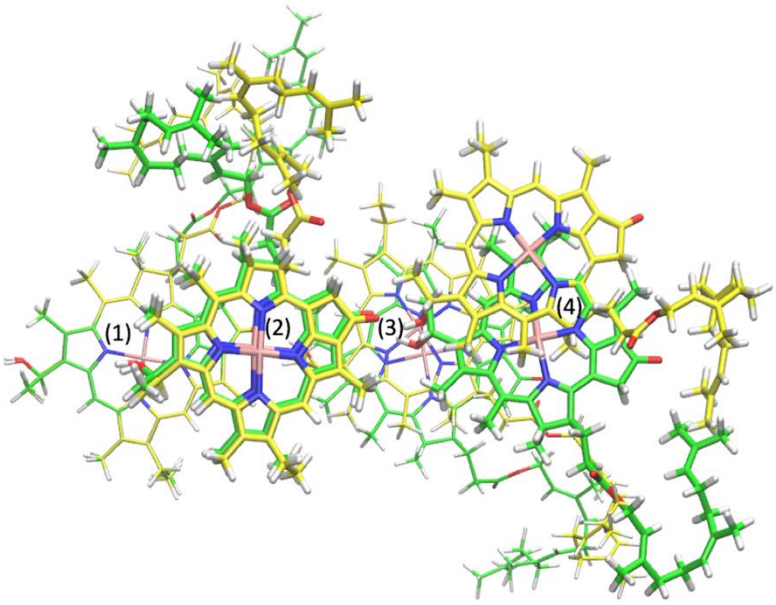

Figure 15: Tetramer structures from structural motif αSβA-dd, displayed in green and yellow for tetramer located in flat and curved section respectively. Numerating each monomer pair (1-4). The superposition was based on monomer (1).

Fig. 15 shows that the local difference between flat and curved geometry is clearly more significant for monomer pairs (3) and (4), while the difference is much smaller in the case of monomer 2 (obviously the difference between both geometries for monomer (1) was negligible since this monomer was used as a reference for RMSD fit). This indicates that the curvature is realized by local displacement of BChl units (monomers) bonded by the hydrogen bond rather than bonding by the stacking interaction. The difference between flat and curved geometry is mostly noticeable for monomer (4).

Since our results of local interaction energy of the BChl tetramers indicated that the local BChl tetramer from the flat structure is more stable than the curved one, then the spontaneous formation of the curvature cannot be explained from the point of view of the local energy. One of the other reasons for chlorosomes to form a curved geometry might be due to entropy. When we consider the tetramer located in the curved section, we can presume it has more degrees of freedom (angles and positions to locate itself and move within the system) compared to one located in the flat section, therefore having larger entropy.

If we want to describe the macrosystem, therefore we find larger abundancy of possible microsystems defining the curved section compared to the case of the flat section. For this, then the entropy of the system would increase when the curvature is formed. According to the second law of thermodynamics, at a certain temperature, systems favour the minimum amount of Gibbs free energy.

$$\Delta G = \Delta H - T\Delta S \qquad (9)$$

ΔG being change of Gibbs free energy, ΔH being the change of enthalpy which represents the system's internal energy, T being the temperature and ΔS is the change of the entropy.

Further studies may consider studying how the system behaves at different temperatures, observing then the effect of entropy on the system geometry which as shown in Eq. (9) are directly related. Therefore, comparing it to the internal energy of the structure in order to identify if entropy is the reason for the curved nature of chlorosomes. Another possible reason for the formation of the curved geometry of BChl layers may be the effect of the total potential energy of the system rather than local interaction energy.

6. Conclusion

To conclude, present work shows evidence of a difference in interaction energy for both plausible motifs (βAβA and αSβA-dd) of the studied tetramer structures located at flat and curved section of curved BChl aggregate. The results indicate, that the BChl tetramer at the flat section is more stable than the one at the curved one from a local energy perspective for both studied structural motifs. Therefore, contradicting the local interaction between BChl molecules as the origin of the formation of the curvature found within the chlorosome layers, caused by the lamellar nature of BChl aggregates, as it has been observed both experimentally and by MD simulations [6, 8, 35-37, 70].

Our calculations also showed weaker interactions between the tetramer and the surrounding in the curved section compared to the flat section. This finding also contradicts the experimental evidence of natural curvature found in chlorosomes.

In addition, the examination and comparison for the effect of curvature on local geometry by superimposing the bacteriochlorin ring of one monomer from the tetramer located in flat section over the same monomer from the tetramer located in curved section showed that curvature is formed by local displacement of BChl units (monomers) bonded by the hydrogen bond rather than BChl units bonded by the stacking interactions.

Seeing that the results of the local energy of BChl tetramers cannot justify the spontaneous formation of the curvature from a local energy perspective in chlorosomes, alternative reasons for curvature formation must be found. Considering the tetramer found in the curved section has more degrees of freedom (therefore having larger entropy), then one of the possibilities explaining the curvature may be the entropy increase when the curvature is formed. Another possible reason, for the formation of a curved geometry of BChl layers, may be the effect of the total potential energy of the system rather than local interaction energy.

7. Literature

[1] Mathias O. Senge, Kevin M. Smith. Biosynthesis and Structures of the Bacteriochlorophylls. In: Blankenship Anoxygenic Photosynthetic bacteria pp 137-151 (1995)

[2] Cohen-bazire *et al*. The fine structure of green bacteria. *J. Cell Biol.* 22, 207–225 (1964).

[3] Bryant, D. A. *et al.* Candidatus Chloracidobacterium thermophilum: An Aerobic phototropic acidobacterium. *Science.* 317, 523–526 (2007).

[4] Eisen, J. *et al.* The complete genome sequence of Chlorobium tepidum TLS, a photosynthetic, anaerobic, green-sulfur bacterium. *Proc. Natl. Acad. Sci. U. S. A.* 99, 9509–14 (2002).

[5] Frigaard, N.-U. & Bryant, D. Seeing green bacteria in a new light: genomics-enabled studies of the photosynthetic apparatus in green sulfur bacteria and filamentous anoxygenic phototrophic bacteria. *Arch. Microbiol.* 182, 265–76 (2004).

[6] Psencík, J. *et al*. The Structural Basis of Biological Energy Generation, Advances in Photosynthesis and Respiration (ed. Hohmann-Marriott, M.) *Springer* 77–109 (2014).

[7] Martinez-Planells, A. *et al.* Determination of the topography and biometry of chlorosomes by atomic force microscopy. *Photosynth. Res.* 71, 83–90 (2002).

[8] Oostergetel, G. T. *et al*. The chlorosome: a prototype for efficient light harvesting in photosynthesis. *Photosynth. Res.* 104, 245– 255 (2010).

[9] Bryant, D. & Frigaard, N.-U. Prokaryotic photosynthesis and phototrophy illuminated. *Trends Microbiol.* 14, 488–96 (2006).

[10] Blankenship, R. E. & Matsuura, K. in Light-harvesting antennas in photosynthesis (ed. Green, B. & Parsons, W.) *Springer* 195–217 (2003).

[11] Staehelin, L. *et al*. Visualization of the supramolecular architecture of chlorosomes (Chlorobium type vesicles) in freeze-fractured cells of Chloroflexus aurantiacus. *Arch. Microbiol.* 119, 269–277 (1978).

[12] Staehelin, L. *et al*. Supramolecular organization of chlorosomes (chlorobium vesicles) and of their membrane attachment sites in Chlorobium Limicola. *Biochem. Biophys. Acta* 589, 30–45 (1980).

[13] Sørensen, P. *et al*. Chlorosome lipids from Chlorobium tepidum: characterization and quantification of polar lipids and wax esters. *Photosynth. Res.* 95, 191–6 (2008).

[14] Buck, D. & Struve, W. Tubular exciton models for BChl c antennae in chlorosomes from green photosynthetic bacteria. *Photosynth. Res.* 48, 367–377 (1996).

[15] Blankenship, R. *et al*. in Anoxygenic photosynthetic bacteria (eds. Blankenship, R. E., Madigan, M. & Bauer, C.) *Kluwer Academic Publishers* 399–435 (1995).

[16] Ikonen, T. P. *et al.* X-ray scattering and electron cryomicroscopy study on the effect of carotenoid biosynthesis to the structure of Chlorobium tepidum chlorosomes. *Biophys. J.* 93, 620–628 (2007).

[17] Arellano, J. B. *et al.* Hexanol-induced order-disorder transitions in lamellar self-assembling aggregates of bacteriochlorophyll c in Chlorobium tepidum chlorosomes. *Langmuir* 24, 2035–2041 (2008).

[18] Arellano, J. *et al*. Nanosecond laser photolysis studies of chlorosomes and artificial aggregates containing bacteriochlorophyll e: evidence for the proximity of carotenoids and bacteriochlorophyll a in chlorosomes from Chlorobium phaeobacteroides strain CL1401. *Photochem. Photobiol.* 72, 669–675 (2000).

[19] Olson, J. & Pedersen, J. Bacteriochlorophyll c monomers, dimers, and higher aggregates in dichloromethane, chloroform, and carbon tetrachloride. *Photosynth. Res.* 25, 25–37 (1990).

[20] Deda, Daiana K., & Araki, Koiti. Nanotechnology, Light and Chemical Action: An Effective Combination to Kill Cancer Cells. *Journal of the Brazilian Chemical Society*, *26* (12), 2448-2470 (2015).

[21] Tamiaki, H. *et al.* In vitro synthesis and characterization of bacteriochlorophyll-f and its absence in bacteriochlorophyll-e producing organisms. *Photosynth. Res.* 107, 133–138 (2011).

[22] Scheer, H. in Chlorophylls and bacteriochlorophylls (eds. Grimm, B., Porra, R., Rudiger, W. & Scheer, H.) *Springer* 1–26 (2006).

[23] Robert E. *et al*. Identification of a Key Step in the Biosynthetic Pathway of Bacteriochlorophyll *c* and Its Implications for Other Known and Unknown Green Sulfur. *Journal of Bacteriology*, 186 (16) 5187-5188 (2004).

[24] T. Jochum. *et al*. The supramolecular organization of self-assembling chlorosomal bacteriochlorophyll c, d, or e mimics, *Proc. Natl. Acad. Sci.* 105 (2008) 12736–12741.

[25] T. Nozawa. *et al*. Structures of chlorosomes and aggregated BChl c in chlorobium tepidum from solid-state high-resolution CP/MAS 13C NMR, *Photosynth.* Res. 41 211–223 (1994).

[26] A. Egawa. *et al*. Structure of the light-harvesting bacteriochlorophyll c assembly in chlorosomes from chlorobium limicola determined by solid-state NMR, *Proc. Natl. Acad. Sci*.104 790–795 (2007).

[27] S. Ganapathy. *et al.* Alternating syn-anti bacteriochlorophylls form concentric helical nanotubes in chlorosomes, *Proc. Natl. Acad. Sci.* 106 8525–8530 (2009).

[28] A.R. Holzwarth & K. Schaffner, On the structure of bacteriochlorophyll molecular aggregates in the chlorosomes of green bacteria. *Photosynth*. Res. 41 225– 233 (1994).

[29] T. Mizoguchi. *et al*. Structural transformation among the aggregate forms of bacteriochlorophyll c as determined by electronic absorption and NMR spectroscopies: dependence on the stereoisomeric configuration and on the bulkiness of the 8-C side chain, *Photochem. Photobiol*. 71 596–609 (2000).

[30] B.J. van Rossum. *et al*. A refined model of the chlorosomal antennae of the green bacterium chlorobium tepidum from proton chemical shift constraints obtained with high-field 2-D and 3-D MAS NMR dipolar correlation spectroscopy, *Biochemistry* 40 1587–1595 (2001).

[31] Y. Kakitani. *et al*. Assembly of a mixture of isomeric BChl c from chlorobium limicola as determined by intermolecular C-13-C-13 dipolar correlations: coexistence of dimer-based and pseudo-monomer-based stackings, *Biochemistry* 45 7574–7585 (2006).

[32] Y. Kakitani. *et al.* Stacking of bacteriochlorophyll c macrocycles in chlorosome from chlorobium limicola as revealed by intermolecular C-13 magnetic-dipole correlation, X-ray diffraction, and quadrupole coupling in Mg-25 NMR, *Biochemistry* 48 74–86 (2009).

[33] Klinger, P. *et al*. Effect of Carotenoids and Monogalactosyl Diglyceride on Bacteriochlorophyll c Aggregates in Aqueous Buffer: Implications for the Self-assembly of Chlorosomes. *Photochem. Photobiol.* 80**,** 572– 578 (2004).

[34] Alster, J. *et al*. Computational study of short-range interactions in bacteriochlorophyll aggregates. *Comput. Theor. Chem*. 998, 87–97 (2012).

[35] Psencík, J. *et al.* Lamellar organization of pigments in chlorosomes, the light harvesting complexes of green photosynthetic bacteria. *Biophys. J.* 87, 1165–72 (2004).

[36] Oostergetel, G. T. *et al.* Long-range organization of bacteriochlorophyll in chlorosomes of Chlorobium tepidum investigated by cryo-electron microscopy. *FEBS Lett.* 581, 5435–5439 (2007).

[37] McDonell, A. Engineering of bacteriochlorophyll pigments and quantum dots for the production of novel photonic materials (2016).

[38] Grant, G. H.; Richards, W. G. Computational chemistry*; Oxford University Press*: Oxford [England]; New York, (1995).

[39] Leach, A. Molecular modeling: principles and applications; 2nd ed.; *Prentice Hall*: Harlow England; New York, (2001).

[40] Jensen, F. Introduction to computational chemistry; Wiley: Chichester; New York, 1, (1999).

[41] McCammon, *et al.* Dynamics of folded proteins. *Nature* 267, 585–90 (1977).

[42] Schlick, T. *et al.* Biomolecular modeling and simulation: a field coming of age. *Quarterly Reviews of Biophysics* 44, 119-228 (2011).

[43] Encyclopedia of computational chemistry 1, A - D.; Wiley: Chichester; Weinheim, (1998).

[44] Schrödinger, E. An Undulatory Theory of the Mechanics of Atoms and Molecules. *Physical Review* 28, 1049–1070 (1926).

[45] Born, M.; Oppenheimer, R. Zur Quantentheorie der Molekeln. Annalen der Physik, 389, 457–484, (1927).

[46] Young, D. C. Computational Chemistry; John Wiley & Sons, Inc.: New York, USA, (2001).

[47] Murphy, R. B. *et al*. A mixed quantum mechanics/molecular mechanics (QM/MM) method for large-scale modeling of chemistry in protein environments. *J. Comp. Chem*. 21, 1442-1457 (2000).

[48] Senn, M. H. et al. QM/MM Methods for Biomolecular Systems. *Angewandte Chemie International Edition*. Vol 48, (2009).

[49] Wiley: Weinheim. Molecular modeling basic principles and applications; (2003).

[50] Slater, J. The Theory of Complex Spectra. *Physical Review*, 34, 1293–1322 (1929).

[51] Becke. D. A. A new mixing of Hartree Fock and local density-functional theories. *J. Chem. Phys* , 98, 1372, (1993).

[52] Stewart, J. J. P. Optimization of parameters for semiempirical methods I. Method. J. *Comput. Chem*. 10, 209–220, (1989).

[53] Dewar, M. J. S.; Thiel, W. Ground states of molecules. 38. The MNDO method. Approximations and parameters. *Journal of the American Chemical Society*, 99, 4899–4907 (1977).

[54] Dewar, M. *et al.* Development and use of quantum mechanical molecular models. 76. AM1: a new general-purpose quantum mechanical molecular model. *Journal of the American Chemical Society*, 107, 3902–3909 (1985).

[55] Hohenberg, P. Inhomogeneous Electron Gas. *Physical Review*, 136, B864–B871 (1964).

[56] S. Grimme, "Semiempirical GGA-type density functional constructed with a long-range dispersion correction," *J. Comp. Chem.*, 27,1787-1799 (2006).

[57] McQuarrie Quantum chemistry, Mill Valey, California: University Science book (2008).

[58] A. Schaefer, H. Horn, and R. Ahlrichs, "Fully optimized contracted Gaussian-basis sets for atoms Li to Kr," *J. Chem. Phys.*, 97 2571-77 (1992).

[59] S.F. Boys, F. Bernardi, Calculation of small molecular interactions by differences of separate total energies – some procedures with reduced errors, *Mol. Phys.* 19 553–566 (1970).

[60] Elspeth F. Garman, Developments in X-ray Crystallographic Structure Determination of Biological Macromolecules, *Science* Vol. 343, 1102-1108 (2014).

[61] Gerhard Wider, Structure Determination of Biological Macromolecules in Solution Using Nuclear Magnetic Resonance Spectroscopy, *Future science* (2018).

[62] Rafael Fernandez-Leiro *et al*, Unravelling biological macromolecules with cryo-electron microscopy, *Nature* 537, 339-346 (2016).

[63] Cramer, C. Essentials of computational chemistry: theories and models; 2nd ed.; Wiley: Chichester West Sussex England; Hoboken NJ, (2004).

[64] Balaban, T. S. *et al.* Relevance of the diastereotopic ligation of magnesium atoms of chlorophylls in Photosystem I. *Photosynth*. Res. 86, 197–207 (2002).

[65] Small-Molecule Drug Discovery Suite 2015-1: QSite, version 6.6, Schrödinger, LLC, New York, NY, (2015).

[66] Banks, J. L. *et a*l. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J. Comp. Chem*. 26, 1752 (2005).

[67] Frisch, M. J. et al. Gaussian 09, revision A.1; Gaussian, Inc., Wallingford, CT, (2009).

[68] Ilya, K. Calculation of Intermolecular Interactions, in Intermolecular Interactions: Physical Picture, Computational Methods and Model Potentials; John Wiley & Sons, Ltd: Chichester, UK, (2006).

[69] Jeziorski, B. *et al*. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van Der Waals Complexes. *Chem. Rev*. 94, 1887-1930 (1994).

[70] J. Psencik. *et al*. The lamellar spacing in self-assembling bacteriochlorophyll aggregates is proportional to the length of the esterifying alcohol, *Photosynth*. *Res.* 104, 211–219 (2010).