Density-functional theory study of gramicidin A ion channel geometry and electronic properties

Milica Todorović1, †, David R. Bowler2,3,4, Michael J. Gillan3,4 and Tsuyoshi Miyazaki1

1National Institute for Materials Science, 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan
2International Centre for Materials Nanoarchitectonics (MANA), National Institute for Materials Science, 1-1 Namiki, Tsukuba, Ibaraki 305-0044, Japan
3London Centre for Nanotechnology, UCL, 17–19 Gordon Street, London WC1H 0AH, UK
4Thomas Young Centre, Department of Physics and Astronomy, UCL, Gower Street, London WC1E 6BT, UK

Understanding the mechanisms underlying ion channel function from the atomic-scale requires accurate ab initio modelling as well as careful experiments. Here, we present a density functional theory (DFT) study of the ion channel gramicidin A (gA), whose inner pore conducts only monovalent cations and whose conductance has been shown to depend on the side chains of the amino acids in the channel. We investigate the ground state geometry and electronic properties of the channel in vacuum, focusing on their dependence on the side chains of the amino acids. We find that the side chains affect the ground state geometry, while the electrostatic potential of the pore is independent of the side chains. This study is also in preparation for a full, linear scaling DFT study of gA in a lipid bilayer with surrounding water. We demonstrate that linear scaling DFT methods can accurately model the system with reasonable computational cost. Linear scaling DFT allows ab initio calculations with 10 000–100 000 atoms and beyond, and will be an important new tool for biomolecular simulations.

1. Introduction

Ion channels reside in cell membranes and control the passage of ions, a vital function in regulating cell behaviour. The gramicidin A (gA) ion channel is one of the smallest and simplest ion channel systems [1–3], and was the target of the first antibiotic used in clinical practice. It is formed from a stack of two short beta-helix polypeptides, as shown in figure 1. Each polypeptide contains 15 amino acids in a sequence of alternating left-handed and right-handed types. The latter helicity is rarely found in natural peptides, and owing to the alternating helicity, all peptide residues face outward into the lipid region, forming the unique helix structure with the beta type backbone structure. As is seen in figure 1, the channel structure forms a narrow pore, through which water molecules or ions can permeate in single file.

In spite of its simple structure, gA shows high selectivity towards ion permeation. It is selective for monovalent cations, such as H+, K+, Na+ or Cs+, showing no measurable permeability for anions or polyvalent cations [4]. Among many ion channels, gA was one of the first whose atomic resolution structure was available. Many experiments have been performed to investigate its transport properties and to clarify the mechanism of its high selectivity. Considering the existence of significant experimental information as well as its simple structure, the gA system is of great importance as a model membrane protein [3].

In the computational studies of the gA, two high-resolution structures have been commonly used so far. They are labelled as 1MAG and 1JNO in the PDB database, respectively. The former is obtained from solid-state NMR for the channel in DMPC membrane [5], whereas the latter one is from solution NMR, in detergent micelles [6]. While these two structures are topologically similar,
they exhibit some variation in helix properties and residue tryptophan 9 (Trp9) orientation (figure 2a). The 1MAG structure helix has $R = 6.5$ residues per helix turn and $d = 4.95$ Å helix turn height (pitch) in contrast to $R = 6.3$ and $d = 4.79$ Å of the 1JNO structure. The structural disparities could have originated from any aspect of experimental procedure, with little to indicate the optimal channel geometry: a classical potential MD study comparing the two structures in a lipid membrane environment predicted that the two backbones after dynamic relaxation starting from two structures become almost the same and that Trp9 spends 80% of the time in the 1JNO orientation and 20% in the 1MAG orientation [7]. Although we do not treat the ion permeation directly in this study, it is important to note the experimental result that the rate of the ion permeation is significantly changed by the replacement of the tryptophan (Trp) residues, which have dipole moments (approx. 2 D), with other residues [6,8–10]. When it is replaced with non-polar phenylalanine, the ion permeation rate is greatly reduced. The change of the ion permeation rate is larger if the number of substituted Trp groups is larger, and also strongly depends on which Trp groups are replaced; the mutation of Trp9 is reported to show the largest change.

The system is an ideal target for theoretical studies. Although most theoretical studies were based on semi-microscopic models in the early stage, it is necessary to employ atomic-scale simulations, for instance molecular dynamics (MDs), for quantitative and detailed understanding of the permeation properties [11]. Many MD simulations of the gA ion channel system have been already reported [7,12–16] and nowadays it is common to treat the gA system in the channel environment, that is gA with surrounding lipid bilayers (a simple model for the membrane) and bulk water. Such MD studies have clarified many aspects of the system, but they have a serious problem that the results sometimes depend on the type of the force field used in the calculations [15,19,20]. It is therefore important to apply more accurate methods to the gA system.

Ab initio approaches based on density functional theory (DFT) have been playing important roles in clarifying the physical properties of various kinds of materials, though standard implementations of DFT are limited in the system size which they can model. A simplified gA model system has been modelled by DFT calculations to study proton diffusion in the pore of the gA channel [21]. A recent quantum mechanics/molecular mechanics study on the gA system showed the importance of polarization of the water molecules in the gA pore [17,22]. However, theoretical studies based on DFT are still very limited because of the demanding computational time and scaling with system size. Recent work has developed DFT methods which scale linearly with system size [23], allowing DFT calculations on systems with 10 000–100 000 atoms or more [24]. This brings the possibility to perform DFT calculations on entire biological molecules and their environment. In this work, we calculate the atomic and electronic structures of the isolated gA using the linear scaling DFT code CONQUEST [25,26]. In particular, we investigate the effects of the side chains of the channel on its structural and electronic properties. Although we only treat the isolated gA here, we still obtain important information about the structural stability and electronic structure of the gA molecule itself. We expect that such information will also help to clarify and improve the accuracy of force fields in the future. To our knowledge, it is the first DFT study to treat the whole gA molecule without any simplification. We will report linear scaling DFT calculations on the gA channel with full environment in a future publication.

The rest of the paper is organized as follows. In §2, we briefly explain the computational method used in this work. The results of our calculations on the atomic and electronic structures of the isolated gA molecule are presented in §3. Finally, concluding remarks are given in §4.

**Figure 1.** (a) Structure of gA ion channel in lipid bilayers. The upper and lower areas are bulk water regions. (b) Top (i) and side (ii) views of the gA molecule.

**Figure 2.** (a) Comparison of 1MAG (blue/dark) and 1JNO (red/light) experimental structures. Ab initio structural optimization initial (black/dark) and final (red/light) structures in the (b) 1MAG case and in (c) 1JNO case.
2. Computational method: local orbital and order-\(N\) methods

The calculations in this work are performed using the CONQUEST code, which is designed for very large-scale DFT calculations [24–27]. The method used in the code is based on DFT, with the pseudopotential technique. In this work, we employ the generalized gradient approximation (GGA) by Perdew, Burke and Ernzerhof (PBE) [28] for the exchange–correlation energy.

In CONQUEST, we work with local orbitals \(\{\phi_a(r)\}\), which are centred on the position of each atom \(i\) with the index of orbital \(a\). We represent the Kohn–Sham orbitals \(\Psi_i(r)\) (\(v\): band index) as linear combinations of the local orbitals.

\[
\Psi_i(r) = \sum_{\alpha} c_{\alpha i}^{(v)} \phi_{\alpha i}(r),
\]

These local orbitals, which we call support functions, are themselves expressed by basis set functions. CONQUEST provides two types of basis set functions: B-splines on regular grids for accurate calculations because the basis set is systematically improvable [29]; and pseudo atomic orbital basis set for efficient calculations [30]. Users can choose one of the two basis set functions depending on the purpose. We only use the latter in this study.

Although CONQUEST can solve for the Kohn–Sham orbitals using a conventional \(O(N^3)\) diagonalization technique, the big advantage of using CONQUEST is that it can also employ \(O(N)\) calculations. In the \(O(N)\) calculations, the code uses the density-matrix minimization method proposed by Li, Nunes and Vanderbilt (LNV) [31,32]. As the detailed explanation of the method has already been given in our previous papers, only the main points are explained in the following. In the density-matrix minimization method, we calculate the density matrix instead of the Kohn–Sham orbitals. Formally, the density matrix is defined as

\[
\rho(r,r') = \sum_{i} \tilde{f}_i \Psi_i^*(r) \Psi_i(r').
\]

The total energy within DFT can be expressed by the density matrix \(\rho\) and we optimize it instead of solving Kohn–Sham orbitals. In the LNV method, the matrix \(K\) is expressed as \(K = 3\text{LSL} - 2\text{LSDL}\) to keep the density matrix weakly idempotent. Then, we introduce a cut-off radius \(R_L\) for the off-diagonal elements of the \(L\) matrix in order to use the locality of density matrix. One important advantage of the LNV method is that it is variational with respect to \(R_L\). We can examine the cut-off energy dependence of total energy easily and can evaluate an error from using a finite cut-off of \(R_L\). This convergence behaviour is reported in §3.3 for the DFT calculation of the isolated gA system.

In both diagonalization and \(O(N)\) methods, the code can perform non-self-consistent (NSC) DFT calculations using the Harris–Foulkes energy functional with a charge density constructed from the superposition of atomic charge [33,34]. Although the NSC method is less accurate than usual DFT calculations with the self-consistent field (SCF) method, it enables us to perform efficient structure optimization or MDs. Note that the code can calculate forces that are completely consistent with the NSC method [27,35]. This method may be useful for our future study on gA systems, and its accuracy is also checked in this paper.

3. Results and discussion

3.1. Structural properties

In this section, we report our results on the structural properties of the gA molecule in the vacuum. Here, we perform structure optimization using two different high-resolution structures as initial atomic positions, 1MAG and 1JNO structures (figure 2a). In order to employ efficient structure optimization, we first use the NSC method, then switch to the usual SCF method [35]. In both NSC and SCF methods, all atomic positions are relaxed using the standard conjugate gradient method until the maximum force component becomes smaller than 0.05 eV/Å. Note that the calculations in §§3.1 and 3.2 are done by the standard diagonalization method with the DZP basis set, explained in §2, and GGA–PBE functional [28]. We use only a \(\Gamma\) point for \(k\)-point sampling and the cut-off energy for the charge density grid is 150 Hartree. Periodic boundary is used with a unit cell of 23.3 Å \(\times 23.3\) Å \(\times 35.0\) Å.

Figure 2b,c show the initial and optimized atomic positions starting from 1MAG and 1JNO structures, respectively. Although the original structures are obtained in different environments (in DMPC bilayers or in SDS micelles), we found that the optimized positions are close to the original ones in both cases. These figures also show that 1MAG has larger differences between the initial and final coordinates, especially for the backbone structure. In order to investigate the backbone structure in detail, we calculate the number of residues per helix turn \(R\) and the helix pitch \(d\) following the method in Ketcham et al. [5]. The root mean square deviations (r.m.s.d.) for the atoms of the backbone are also calculated. These structural parameters of 1MAG and 1JNO, before and after the structure optimization, are listed in table 1. The results show that the optimized 1MAG and 1JNO have different backbone structures. In addition, the change in backbone structure is small in 1JNO case, whereas it is significant in 1MAG case, especially for helix pitch \(d\). Although the original value of \(d\) in 1MAG case is already larger than that of 1JNO, the increase of \(d\) in 1MAG case is 0.15 Å, which is larger than 0.08 Å in 1JNO case. We observe a similar behaviour in the hydrogen bond lengths. The \(\beta\)-helix structure is stabilized by the hydrogen bonds formed between the backbone amide hydrogen and the carbonyl oxygen atoms parallel to the pore axis. The increase of the hydrogen bonds by the structure optimization is 12\% in 1MAG, whereas it is 3\% in 1JNO case.

For the 1MAG structure, it is useful to know what causes the large change of the backbone structure during the optimization: either the effect of the backbone itself or that of the side chains. One of the main differences between 1MAG and 1JNO structures is the different orientation of the indole ring of one tryptophan (Trp9). The atomic positions of other side chains are also slightly different. These differences may cause the present different backbone structures. In order to investigate the effects of side chains on the backbone structure, we prepare a fictitious channel by replacing all amino acids of the gA with alanine, keeping the same backbone structure as the original 1MAG structure, shown in figure 3. Hereafter, we refer to this system as an alanine tube. We have relaxed the atomic positions of this system and analyse the values \(R\) and \(d\) of the resulting optimized structure. As is shown in table 1, the change of the structure by the relaxation is found to be very small. The r.m.s.d. value is 0.21 Å and the change of \(d\) is only 0.04 Å. This result
suggests that the large change in helix pitch during the structure optimization of the 1MAG structure is owing to the repulsion interactions between the side chains in the original 1MAG structure.

From the results so far, we can conclude that the original 1JNO structure is close to the optimized geometry in the gas phase, while 1MAG needs a large relaxation to reach the stable structure. By comparing the total energy, we find that the initial 1JNO structure is more stable than 1MAG by 10.3 eV. The corresponding energy difference calculated by classical force fields is 10.4 eV by CHARMM (v. 27) and 19.7 eV by AMBER (v. 99). The value found using AMBER is larger than the DFT value, but these values are essentially consistent with the DFT result. If the 1MAG structure is less stable than 1JNO by such a large amount of energy, it is unlikely that the 1MAG structure will exist even in a lipid bilayer. However, after structure optimization with DFT, this large energy difference between the two structures is greatly reduced. The final structure of 1JNO is still more stable than 1MAG, but only by 1.8 eV. Although this energy difference is not negligible, it is small enough that the relative stability may change when we introduce the interactions between lipids and gA or include entropic effects, which are often important in soft biological molecules. If the 1MAG structure gains more energy from these effects than 1JNO, then the 1MAG structure might be more stable.

We now report the stabilization energy of the two helices of gA. When the gA ion channel is open for ions, the upper and lower helices must be associated to form a dimer. They are also sometimes dissociated by the lateral movement in the membrane (lipid bilayers) and in such cases the channel is closed for ion permeation. Thus, the stabilization energy of the two helices can be regarded as a measure for the stability of the open channel structure. In this respect, the stabilization energy is important and we calculate the energy defined as

\[ E_a = E_{\text{dimer}} - 2E_{\text{mono}}. \]  

Here, \( E_{\text{mono}} \) is the total energy of the upper or lower peptides in isolation, whose atomic positions were relaxed. \( E_{\text{dimer}} \) is the total energy of the associated gA, calculated with the basis set superposition error correction using the counterpoise method [36]. With the optimized 1JNO structure, \( E_s \) is calculated to be 0.88 eV. This value is reasonable given the number of hydrogen bonds between the two helices, which have three N-H-O and three O-H-N bonds. Note that the exchange–correlation functional of PBE has been reported to be accurate for the expression of hydrogen bonds [37,38]. In the membrane, this associated gA dimer structure will also be affected by the movement of surrounding lipid molecules. It will be important to compare this stabilization energy with lipid–gA and lipid–lipid interactions in the future.

Finally, we note that, while we have mainly discussed results obtained with the SCF method, the differences between the NSC and SCF methods are small for the optimized structures (table 1). As the NSC method is stable, and more efficient, it may play important roles in the future study of larger and more complicated full gA systems.

### 3.2. Electronic structure

In this section, we are particularly interested in the effects of the side chains of the ion channel on the electronic structure in the pore region of the gA. As mentioned in §1, it is known that...
that the substitution of Trp significantly affects the rate of ion permeation [6,8–10]. It is important to understand whether such differences come from the different electronic structures of the side chains or some other factors, for example, from the structural effects of the interactions of the side chains with lipid molecules. We note that there is a report showing that the conductance of a monovalent cation is correlated with the electrostatic interaction energy between the ion and the dipole moment of the indole ring of the tryptophans [9].

First, we show in figure 4 the density of states of gA of 1MAG optimized structure (solid line), as well as that of alanine tube (dotted line) introduced in the last section. We also calculated the DOS of the 1JNO structure, but the difference to the 1MAG DOS is small, particularly around the band gap, so this is not shown for clarity. We can see that the density of states of the 1MAG shows a HOMO–LUMO energy gap of 3.3 eV. The energy gap of alanine tube is much larger than that of gA, at around 4.2 eV, implying that HOMO and/or LUMO of gA mainly come from its side chains. The alanine tube also shows a very low DOS around 4 eV below the Fermi level, which is not present in the gA DOS.

We studied the effects of side chains on the pore of the gA by investigating the charge density and the electrostatic potential. Ion transport depends critically on the electrostatic potential inside the channel pore, which in turn depends strongly on the charge density. The potential also affects the dynamics of the water molecules arranged in the pore. If the charge density is plotted as a function of distance from the channel axis, i.e. the centre of the pore, the two systems are found to be almost identical until the distance is greater than 4.4 Å, the approximate radius of the gA. We can thus conclude that the effects of the side chains on the charge distribution in the pore region are very small.

We next investigate the electrostatic potential, calculated as a sum of the local pseudopotential and the Hartree potential using the self-consistent charge density. Figure 5a shows a cross section of the potential perpendicular to the channel axis for the 1MAG structure, 9.6 Å along the axis away from the centre of the gA. The binding site for the TI (thallium) ion, which is similar to K⁺ ion in size and hydration properties, is reported to be at this position [39]. The calculated binding site using the 1MAG structure is also reported to be close to this position, 9.7 Å [19]. In the figure, areas of lower potential are located near the peptide bonds which form hydrogen bonds. By contrast, the region near the centre of the pore shows a high but flat potential, making it accessible to positively charged particles. Figure 5b shows the equivalent cross section for the alanine tube having a 1MAG backbone structure. There is a clear difference between figure 5a,b in the region of side chains, outside of the backbone. However, the two potentials are almost the same in the pore region. Figure 5c shows the difference of these two potential surfaces, implying that the two pores exhibit identical electrostatic potential at least within 1 meV. The same behaviour can be seen in figure 5d,e, which are cross sections in other planes, specifically the zz and yz planes, showing the potential difference along the channel axis (z-axis). We can conclude from these results that the effects of dipole moments of Trp groups are well screened and the electronic structure of the side chains of the gA does not affect the electrostatic potential in its pore. In order to explain the effect of Trp groups on the ion permeation rate, dynamical effects must be considered. We suggest that the side chains of Trp groups may play important roles in the dynamics of the gA channel through the interactions with water molecules and lipids, especially with their polar heads. This aspect should be clarified in future work.

**3.3. Order-N calculations on the isolated gramicidin A**

We expect that the effects of the membrane or water molecules will be clarified by the all-atom DFT simulations of the gA channel in a lipid membrane with bulk water and counterions as the environment. We are now working to perform large-scale DFT simulations of the full gA system and the results will be reported in a future publication. Once we can calculate the total energy and forces acting on atoms accurately with DFT for many snapshots of the full gA systems, and use the results with the so-called force matching method [40], it will be possible to generate a DFT-derived force field, with which we can investigate dynamics of the biological system. To perform such large-scale DFT calculations, we will need to use an O(N) method. As a preparation for these future studies, we have made preliminary O(N) calculations on the isolated gA, in order to clarify the accuracy of the method on the gA systems. In particular, we can test the effect of the truncation of the density matrix because we can perform exact calculations (we have similarly tested convergence on a DNA ten-mer in water [38] which found essentially complete convergence for a range of 10 Å). Here, we use the NSC method because the convergence behaviour will be qualitatively the same for the SCF methods. For the basis sets, contracted DZP basis sets [30] are used.

Figure 6 shows the total energy calculated for different values of the cut-off, $R_C$. We find that all calculations are stable in the optimization of L matrix. As the present system contains only 552 atoms, we can employ the exact diagonalization method and the total energy calculated by the method is also plotted by a dotted line in the figure. The inner panel of figure 6 shows the difference between the O(N) and exact diagonalization methods, on a log scale. This shows that the total energy by the O(N) method converges very fast towards the exact value. When the cut-off larger than 7.5 Å is used, the energy discrepancy is less than 1 meV per atom. This cut-off is small enough to realize actual calculations. In our previous O(N) DFT study of a different system, Ge nano structures on Si(001) surface, we actually used 10.8 Å for the structure optimization of the systems containing more than 20 000 atoms [41].
We expect that similar accuracy in the total energy would be obtained in the calculations of full gA systems, because the water and lipid molecules have large HOMO–LUMO energy gaps, resulting in fast convergence of the total energy with respect to $R_L$. In fact, in our investigations of dry and hydrated DNA systems, we found that the difference caused by bulk water is extremely small [38]. Based on these results, we expect that the O($N$) calculations of the full gA systems will be both highly accurate and useful for understanding structure and function. These studies are already underway and we have recently succeeded in employing a self-consistent DFT calculation on a full gA system, which consists of a gA channel embedded in 64 DMPC molecules and the water region made of 2473 water molecules.

**Figure 5.** DFT electrostatic potential (eV): (a)–(c) are cross sections perpendicular to the channel axis (z) at 9.6 Å from the centre of the channel. (a) 1MAG geometry; (b) alanine channel; (c) 1MAG and alanine channel potential difference. Note the different scale in (c). Panels (d) and (e) are also potential difference between 1MAG and alanine channel, but cross sections along the channel axis, xz and yz planes, respectively. The contours in (c)–(e) are drawn at $10^{-1}$, $10^{-2}$, $10^{-3}$, $10^{-4}$ and $10^{-5}$ eV for both positive and negative values.

**Figure 6.** Order-$N$ band energy behaviour with increasing range of auxiliary density matrix $L$ (inset: convergence to diagonalization result). (Online version in colour.)
We have also investigated the electronic structures of gA and the alanine tube. The charge density and electrostatic potential in the pore of the two systems have been compared to clarify the effects of side chains. The potential is relevant for the ion transport and the dynamics of the water molecules arranged in single file in the channel. We found that the potential in the pore is almost identical between the two systems. The results imply that the sensitivity of the channel function to the type of residues cannot be explained by the electrostatic interactions made by the dipole of Trp groups. This does not support some of the proposals in previous reports.

In the next step, it is desirable to simulate the system in the channel environment, that is gA in membrane with bulk water. We have already started to work on this large and complex gA system containing more than 16 000 atoms using an O(N) DFT method. The code CONQUEST used in this work has the ability to treat very large systems containing many thousands of atoms, with the density-matrix minimization method. As a preparation for such large-scale DFT studies, we have applied the method to the isolated gA in this paper. By calculating the total energy dependence on the cut-off of the auxiliary density matrix $R_c$, we have found that the total energy converges very fast to the exact value. The errors from using the finite cut-off of $R_c$ become negligibly small if we use $R_c$ larger than 10 Å. It is encouraging and we believe that the method will become one of the standard theoretical techniques to study the atomic and electronic structures of complex biological systems.

Acknowledgements. The authors gratefully acknowledge useful discussions with Antonio Torralba, Michiaki Arita, Takao Otsuka and Wataru Shinoda. Most of the calculations in this work were done using the numerical materials simulator at the National Institute for Materials Science (NIMS), Japan.

Funding statement. This work was partly supported by the Royal Society, Grant-in-Aid for Scientific Research on Innovative Areas (No. 22104005), the Strategic Programs for Innovative Research (SPIRE), MEXT and the Computational Materials Science Initiative (CMSI), Japan.

Endnote

1 A much larger difference between 1MAG and 1JNO is reported in Allen et al. [7].

References

10. Andersen OS, Greathouse DV, Providence LL, Becker MD, Koppel RE. 1998 Importance of tryptophan
dipoles for protein function: 5-fluorination of tryptophans in gramicidin A channels. J. Am. Chem. Soc. 120, 5142 – 5146. (doi:10.1021/ja980182i)


