




Cite this: *Org. Biomol. Chem.*, 2022, **20**, 1306

Conformational preferences induced by cyclization in orbitides: a vibrational CD study†

Maria A. S. Yokomichi,^a Hanyeny R. L. Silva,^b Lorenza E. V. N. Brandao,^c Eduardo F. Vicente^c and Joao M. Batista Jr. *^a

Orbitides are bioactive head-to-tail natural cyclic peptides from plant species. Their bioactivity is intrinsically related to the main conformations adopted in solution, whose correct characterization represents an important bottleneck for medicinal chemistry applications. To date, NMR spectroscopy has been the most frequently used technique to assess the secondary structure of orbitides. Despite the amount of structural information commonly available from NMR, its time scale frequently results in a limited conformational ensemble with a single mean structure, which may not represent the bioactive conformation. Additionally, problems with inter-residue NOE/ROE signals can reduce the accuracy and confidence of the 3D assignments. Vibrational circular dichroism (VCD), on the other hand, has been demonstrated as a powerful tool to probe the stereostructure of chiral molecules, including peptides and proteins, with enhanced sensitivity to individual conformations in the condensed phase. Herein, we present the first VCD stereochemical investigation of orbitides. By combining IR/VCD experiments in ACN-*d*₃ and ACN-*d*₃/D₂O mixtures with DFT calculations in different levels of theory we were able to determine the solution-state conformational behavior, as well as the main structural restraints induced by cyclization, of the seven-residue orbitide pohlianin A and its linear precursor. VCD results indicated inverse γ -turns as the most prevalent structural motif for the linear precursor in partially aqueous solution, while type I and type VI β -turns were induced by the cyclization process, along with some classic γ -turns. In addition to the conformation of the cyclic peptide already described from NMR data, two previously unidentified conformations with distinct secondary structures were found to significantly populate the sample. This conformational discriminatory power of VCD for both linear and cyclic peptides in different solvent systems may lead to more accurate structural characterization of turn-rich peptidic natural products and help the design of conformationally-tailored peptides for more precise structure–activity relationship results.

Received 4th November 2021
Accepted 18th January 2022

DOI: 10.1039/d1ob02170b

rsc.li/obc

Introduction

Orbitides comprise a class of ribossomally synthesized and post-translationally modified head-to-tail (N-to-C) cyclic peptides from higher plant species.¹ These peptides range generally from 5 to 12 L-configured amino acid residues and lack disulfide bonds. Recently, a 16-residue orbitide was reported from *Ratibida columnifera*.² Orbitides have been described from a limited number of families including Annonaceae,

Caryophyllaceae, Euphorbiaceae, Lamiaceae, Phytolaccaceae, Rutaceae, Schisandraceae, and Verbenaceae.³ Despite being restricted to a few families, to date, 191 orbitide sequences have been reported in the dedicated database of cyclic proteins CyBase (<http://www.cybase.org.au>), which demonstrates the potential structural variability of this class of compounds. Orbitides, and cyclic peptides in general,⁴ are considered privileged structures in medicinal chemistry due to higher target affinity, resulting from limited size and flexibility, as well as enhanced membrane permeability and resistance to protease cleavage.⁵ As a result, a number of biological activities have been reported for orbitides, such as cytotoxic,⁶ antimalarial,⁷ immunosuppressive,⁸ antibacterial⁹ and antifungal.¹⁰ Since orbitides are isolated in small amounts, peptide synthesis protocols that include the assembling and cyclization of linear precursors are normally employed. Interestingly, some studies have demonstrated orbitides to be more active than their linear analogues, as in the case of the antimalarial potential of ribifolin.¹¹ Nevertheless, the linear form of this same orbitide

^aInstitute of Science and Technology, Federal University of Sao Paulo (Unifesp), Rua Talim 330, Sao Jose dos Campos, SP 12231-280, Brazil.

E-mail: batista.junior@unifesp.br

^bSao Paulo State University (Unesp), School of Agricultural and Veterinarian Sciences, Via de Acesso Professor Paulo Donato Castellane Castellane S/N, Jaboticabal, SP 14884-900, Brazil

^cSao Paulo State University (Unesp), School of Sciences and Engineering, Rua Domingos da Costa Lopes, 780, Tupa, SP 17602-496, Brazil

† Electronic supplementary information (ESI) available: MS and VCD spectra, dihedral angles, atomic coordinates. See DOI: 10.1039/d1ob02170b

was found to be substantially more potent than the cyclic compound in a fibroblast migration assay.¹² The linear analogues of the orbitides pohlianins A and B were also more active than the cyclic peptides in the same migration assay. The authors have attributed the difference in bioactivity between cyclic and linear peptides, among other features, to their 3D-structures and conformational dynamics in solution. Therefore, correctly assessing the conformational preferences of orbitides and their linear analogues in solution is of paramount importance to understand their bioactivity. Since linear and cyclic peptides frequently adopt multiple conformations in solution, which are difficult to resolve using NMR,¹³ their configurational and conformational studies represent an important bottleneck for medicinal chemistry applications. It becomes even more critical as the number of residues increases. Due to the size and structural restrictions of orbitides, secondary structure motifs such as helices or sheets, where coupling of repeating units overcomes effects of local deviations to give reproducible band shapes, are not observed. On the other hand, orbitides are prone to adopt turns, which are highly variable and local, making them not easily predictable and detectable.

The amino acid absolute configurations of orbitides are generally determined by GC or HPLC analysis of acid hydrolysates, while conformations and secondary structures are assessed predominantly by NMR spectroscopy, sometimes assisted by distance geometry (DG) and molecular dynamics (MD) simulations. X-ray crystallography has been used only in few cases due to the challenging crystallization of some orbitides. Racemic crystallographic procedures have recently been proposed as an alternative to generate orbitide crystal structures of sufficient quality.¹⁴ Although NMR has been used as the technique of choice for orbitide analysis, in the case of flexible molecules in solution, the time scale of NMR spectroscopy may prevent the assessment of fast interconverting conformers, resulting in a limited conformational ensemble with a single mean structure. Additionally, the low number of inter-residue NOE/ROE signals commonly observed for orbitides reduce the accuracy and confidence of their 3D assignments.¹⁴ Vibrational circular dichroism (VCD) spectroscopy, on the other hand, has been demonstrated as a highly sensitive stereochemical tool to determine solution-state secondary structures of peptides and proteins in conditions not commonly accessible to other techniques.¹⁵ Based on the fast time scale of vibrational (and electronic) transitions, despite intrinsic lower resolution compared to NMR, the final IR/VCD spectra contain rich structural information from individual conformers weighted by their Boltzmann factors. Notwithstanding its structural and conformational discriminatory power, including sensitivity to turns and extensive applications to peptides and proteins, VCD has not been used to the study of orbitides before. Regarding other cyclic peptides, VCD has been successfully used to study β - and γ -turn motifs in model peptides in different solvents, with some spectrum-structure relationships proposed.^{16–25} Electronic CD has also been applied to the characterization of β -turns.²⁶ Based on the potential applicability of VCD to assess the conformational

ensembles of both linear and turn-rich cyclic peptides, and the current lack of VCD reports on orbitides, herein, we present a thorough VCD investigation, assisted by MD and density functional theory (DFT) calculations, of synthetic pohlianin A (**1**, Fig. 1).⁷ Pohlianin A is an antimalarial heptapeptide first isolated from *Jatropha pohliana*, in 1999. This orbitide was selected as a model compound since its conformational behaviour has already been described based on NMR data in DMSO solution. Additionally, as the present study uses a synthetic version of pohlianin A, we also had access to its linear precursor (**1a**, Fig. 1). Therefore, the same experimental and theoretical VCD protocols were applied to the linear heptapeptide. It has allowed us to determine its predominant conformations, as well as the conformational restrictions induced by the head-to-tail cyclization step in partially aqueous solution. Finally, we demonstrate the advantages and some of the limitations of using VCD and DFT calculations as a stereochemical tool to determine the 3D structure of orbitides, therefore, paving the way for a more widespread use of this technique by the natural product community. As VCD results in mirror-imaged spectra for enantiomers, not only is conformational analysis feasible, but also absolute configuration assignments. DFT calculations on these large systems have been proved feasible using different combinations of hybrid functionals and basis sets, which resulted in similar spectral features. Further theoretical and experimental VCD investigations of orbitide sequences are underway and may contribute to the design of conformationally-tailored and fine-tuned⁴ cyclic peptides.

Results and discussion

In 2015, Raman optical activity (ROA) was used to assign the absolute configuration and the predominant conformations of

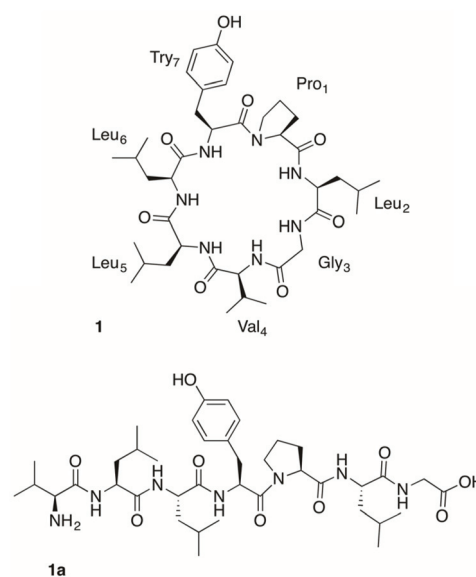


Fig. 1 Chemical structures of the orbitide pohlianin A (**1**) and its linear precursor (**1a**).

ribifolin, an antimalarial orbitide from *Jatropha ribifolia*.¹¹ This ROA study represents the first and still only vibrational optical activity (VOA) investigation of orbitides reported in the literature. The results obtained for the octapeptide ribifolin indicated a single predominant conformer (90% of Boltzmann distribution) in 50 : 50 acetonitrile/water solution, displaying two γ -turns, one being classical (between Leu₂-CO and Ser₄-NH) and one inverse (between Ile₆-CO and Gly₈-NH), which was in accordance with NMR results. The current investigation is focused on the heptapeptide pohlianin A, that differently from ribifolin, contains an aromatic tyrosine residue as well as a proline residue, which plays important roles in the conformational preferences of peptides. In its first report, pohlianin A was confirmed as an all-L peptide and its predominant conformation in DMSO was described as containing a conventional type I β -turn at Val₄-Leu₅ ($i + 1$ and $i + 2$, respectively) stabilized by a hydrogen bond between Leu₆-NH and Gly₃-CO, and a β -bulge (H-bond between Gly₃-NH and Leu₆-CO) motif with a type VIa β -turn around Tyr₇-Pro₁ (H-bond between Leu₂-NH and Leu₆-CO), with *cis*-tyrosyl-proline amide bond, and stabilized by stacking effects. It is important to mention that the numbering presented herein for pohlianin A differs from that of the original publication, since it follows more recent guidelines.³ For the VCD stereochemical investigation of pohlianin A, we performed the synthesis of the all-L and all-D versions of this orbitide (PLGVLLY) by means of solid phase peptide synthesis (SPPS).²⁷ The synthetic approach involved the assembling of a linear precursor (VLLYPLG) following published procedures.¹² The chromatogram profiles and mass spectra of all-L and all-D pohlianin A are shown in the ESI file (Fig. S1 and S2[†]). The IR/VCD spectra of all-L and all-D pohlianin A were recorded in 75 : 25 ACN-*d*₃/D₂O and 100% ACN-*d*₃ solutions at a concentration of 100 mg mL⁻¹. Due to solubility issues in

100% ACN-*d*₃, the experimental spectra of the linear precursor **1a** were measured at the same concentration only in 75 : 25 ACN-*d*₃/D₂O. Measurements at different concentrations ruled out aggregation of the peptides at the measurement conditions. For the calculations, MD conformational analyses were performed for the all-L linear and cyclic forms of pohlianin A starting from different input geometries, and not considering the main conformation derived from NMR data to avoid any bias. MD simulations were run on periodic conditions with explicit water molecules. Other solvents were not used for the simulations due to limitations of the software available in the laboratory. In any case, water made up 50% of the 75 : 25 ACN-*d*₃/D₂O (v/v) solution used in the measurements of **1** and **1a**. MD snapshots were taken at regular intervals and the geometries optimized using DFT at different levels of theory with implicit solvation.

Regarding the experimental IR data, as expected, the spectra obtained for the enantiomers in each enantiomeric pair of cyclic and linear pohlianin A were very similar, with the only notable difference being the amide II region measured in 75 : 25 ACN-*d*₃/D₂O, where discrepancies in band intensities were observed (Fig. 2). It is related either to non-uniform isotopic exchange observed in the presence of D₂O, even after several hours of measurement, or issues arising from solvent subtraction. The VCD spectra of the enantiomers of both linear and cyclic peptides displayed mirror image relationships over the full experimentally accessible wavenumber range (Fig. 2). The forthcoming spectral analysis will be focused on the amide I and amide II regions (denoted I' and II' for N-D), which are the most informative in VCD spectroscopy.

By first comparing the IR and VCD spectra of cyclic pohlianin A (**1**) in 100% ACN-*d*₃ and 75 : 25 ACN-*d*₃/D₂O we can speculate that the cyclic compound adopts similar conformations in both solvent systems (Fig. 2, centre and right and

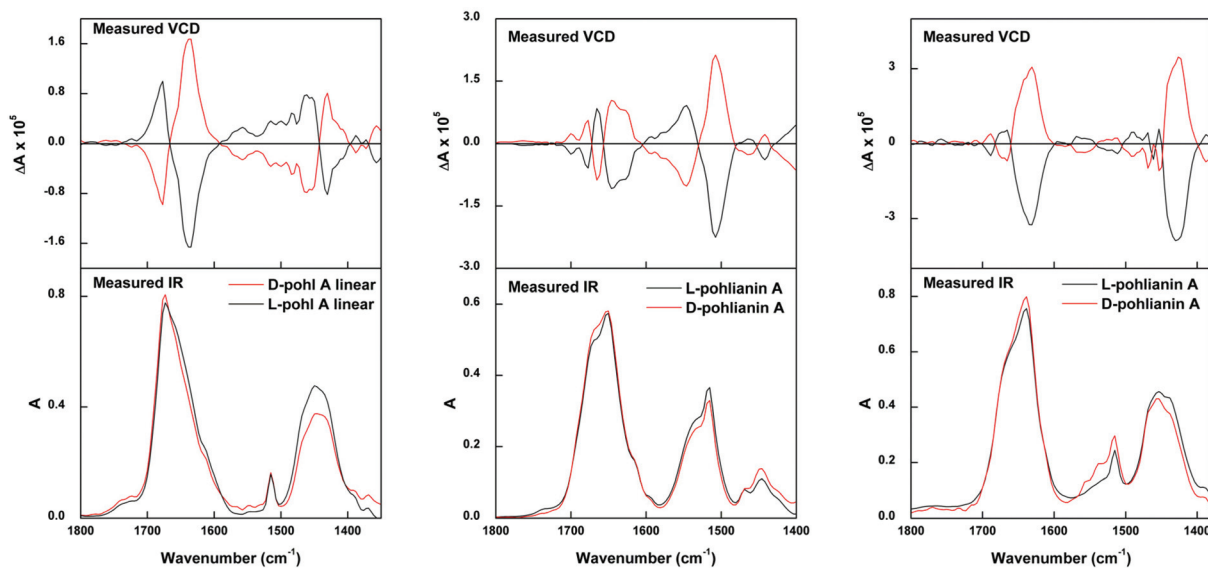


Fig. 2 (Left) Experimental IR and VCD spectra of all-L and all-D **1a** (linear) in 75 : 25 ACN-*d*₃/D₂O; (Centre) experimental IR and VCD spectra of all-L and all-D **1** (cyclic) in 100% ACN-*d*₃; (Right) experimental IR and VCD spectra of all-L and all-D **1** (cyclic) in 75 : 25 ACN-*d*₃/D₂O.

Fig. S3†). The main differences observed in the IR spectra are related to isotopic exchange and the resulting frequency shifts. In the VCD spectra, besides the shift of the amide II' vibrations to lower wavenumbers in ACN- d_3 /D $_2$ O, a broadening of the positive feature at 1546 cm^{-1} (1475 cm^{-1} when deuterated) was observed in partially aqueous solution. In the amide I', the $-,+,-$ feature was retained, even though the relative band intensities changed (Fig. S3†). It is important to mention that the amide I bands of deuterated proteins (amide I') are commonly affected by the NH to ND exchange stemming from NH bending contributions to the amide I (predominantly C=O stretching). Additionally, these variations of the amide I signal arising from deuteration are generally difficult to reproduce by DFT calculations.²⁸ The fact that the $-,+,-$ amide I feature was retained in different solvents and well reproduced by the calculations (Fig. 5 and Fig. S6†) corroborates similar conformations of cyclic pohlianin A in both 100% ACN- d_3 and 75 : 25 ACN- d_3 /D $_2$ O. Based on the higher resolution observed in 100% ACN- d_3 solution, however, further spectral analysis will be focused only on this solvent system for the cyclic peptide.

The experimental IR/VCD spectra of the linear precursor of pohlianin A (**1a**) in 75 : 25 ACN- d_3 /D $_2$ O presents some similarities with the spectra of the cyclic peptide (**1**) in the same solvent system (Fig. 2 left and right and Fig. S4†). The main difference observed in the IR spectra was a blue shift of the maximum absorbance of the amide I' band from 1638 cm^{-1} to 1673 cm^{-1} in the linear precursor. Such a shift may indicate significant environmental changes around the carbonyls reflecting conformational restrictions arising from the cyclization process. Regarding VCD features, the linear peptide resulted in a negative (from low to high wavenumbers) couplet like signal ($+,-$) in the amide I' region centred at 1665 cm^{-1} instead of a $-,+,-$ feature; while the amide II' region presented a higher intensity positive band at 1461 and a lower intensity negative absorption at 1430 cm^{-1} compared to the cyclic peptide. A purely empirical analysis of the amide I' IR/VCD spectra of the linear and cyclic versions of pohlianin A would indicate a PPII conformation for the linear precursor (**1a**) and the presence of some α -helical secondary structure for cyclic pohlianin A (**1**),¹⁵ which would be difficult for such a restricted peptide. In order to correctly interpret the experimental spectra of both forms of pohlianin A, their IR/VCD spectra were simulated using DFT. Comparisons between observed and simulated data were performed for the linear precursor at the B3PW91/PCM(DMSO)/6-31G(d,p) (Fig. 3) and B3PW91/PCM(DMSO)/cc-pVDZ (Fig. S5†) levels. Implicit solvation in DMSO was selected since the 75 : 25 proportion of ACN- d_3 /D $_2$ O represents a water molar fraction of roughly 0.5. The dielectric constant for such a mixture of ACN and water at 25 °C is approximately 46.5,²⁹ which is very close to that of DMSO. Since implicit solvation does not take into account intermolecular hydrogen bonds, we judged the use of the dielectric constant of DMSO more realistic than that of either water or acetonitrile alone. Unfortunately, for such large systems, the quantum chemical consideration of explicit solvent molecules may be prohibitive. Explicit solvation was tentatively con-

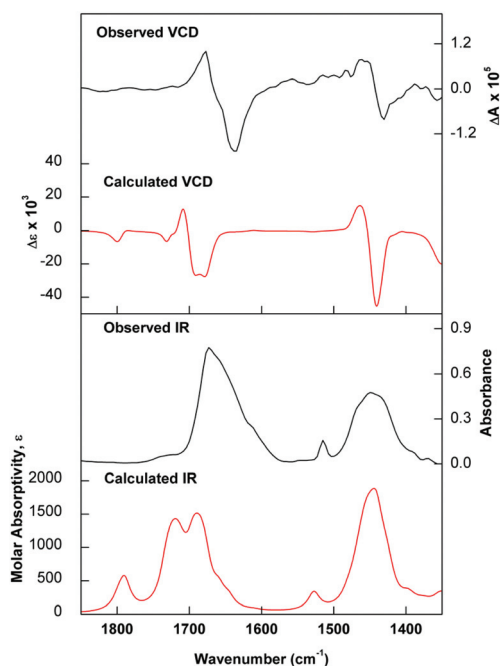


Fig. 3 Comparison of experimental IR and VCD spectra of **1a** (linear) recorded in 75 : 25 proportion of ACN- d_3 /D $_2$ O (black trace) with calculated [B3PW91/PCM(DMSO)/6-31G(d,p)] IR and VCD spectra for the simple average of the four lowest-energy conformers identified for all-**1a** (linear) (red trace).

sidered only by means of the two-layer ONIOM approach. Exchangeable hydrogen atoms attached to oxygen and nitrogen were replaced with deuterium to account for the isotopic exchange observed in the presence of D $_2$ O. Despite the large number of conformers generated by the MD simulations, only four conformers were predicted to significantly populate the sample at the B3PW91/PCM(DMSO)/6-31G(d,p) level at the working temperature. The very good agreement between experiment and DFT-simulated data (Fig. 3) allowed us to determine the main conformations adopted by compound **1a** in 75 : 25 ACN- d_3 /D $_2$ O (Fig. 4). The QM/MM simulations with explicit water afforded very similar geometries for the above-mentioned conformations, even though the simulated spectra were somehow different, lacking the positive VCD amide I' band (Fig. S9†). The overall IR agreement with experiment, as well as that of the amide II VCD band, was improved, though. Another approach tested involved simulations considering zwitterionic forms of the conformers identified for **1a**. The results were similar to those presented in Fig. 3 for the non-ionized molecule (Fig. S10†). Overall, both forms are appropriate to reproduce the experimental spectra of **1a** in 75 : 25 ACN- d_3 /D $_2$ O and explicit consideration of charges in either N or C termini does not seem critical. This is relevant since zwitterionic forms of **1a** presented more frequent convergence issues when compared to the non-ionized counterparts, especially when PCM (DMSO) was included. The four lowest-energy conformers identified (**1a.1**–**1a.4**, Fig. 4) were used as a single average to avoid population bias arising from DFT energy cal-

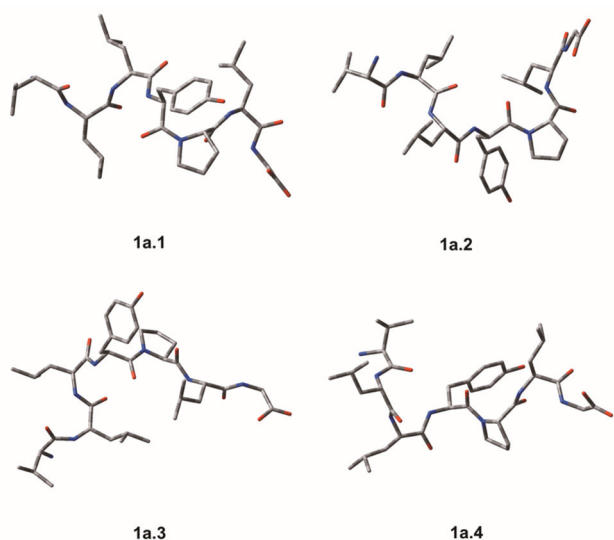


Fig. 4 Lowest-energy conformers of **1a** (linear) identified at the B3PW91/PCM(DMSO)/6-31G(d,p) level and used as a single average. Hydrogen atoms omitted for clarity.

culations (Table S2 and Fig. S7†). The inherent flexibility of the small-sized linear peptide resulted in the lack of long repeating sets of ϕ/ψ angles typical of classical secondary structure motifs among the identified conformers. Consequently, the observed IR and VCD features arose from contributions of local conformations, with short-range coupling. This is particularly important for the amide I' VCD band, whose shape in the experimental spectrum results from individual contributions of different signs being averaged out (Fig. S7†). Although the experimental amide I' band shape suggests the presence of left-handed helix structure, it was not corroborated by the lowest-energy conformers identified, as will be described below. The deuterated amide II region, besides N-D bendings and C-N stretching modes, also contains contributions from CH₂ scissoring and CH₃ deformation modes of the side chains. Conformer **1a.1** presents 2 inverse γ -turns involving Leu₂ and Leu₅ in the $i + 1$ position, while Tyr₇ adopts dihedral angles typical of β -sheet (−156.8, 139.0) and Pro₁ those typical of PPI (−77.9, 154.3). Conformer **1a.2** presents 3 inverse γ -turns with the following residues in the $i + 1$ position: Pro₁, Leu₅, and Leu₆. Tyr₇ adopts dihedral angles typical of β -sheet (−154.5, 139.5). Conformer **1a.3** has a single inverse γ -turn involving Leu₅ in the $i + 1$ position. Residues Pro₁ and Leu₂ present ϕ/ψ angles of (−63.9, 144.6) and (−128.4, 151.1), commonly present in PPII and β -strand forming residues, respectively. Conformer **1a.4** presents a type I β -turn at Leu₅–Leu₆ stabilized by a hydrogen bond between Val₅–CO and Tyr₇–NH. It is important to mention that conformer **1a.1** was the only one to present a *cis* tyrosyl–proline amide bond. All the remaining amide bonds in all conformers had ω angles of approximately 180°. Therefore, instead of left-handed helix structures, the +,− amide I' feature in this linear heptapeptide may indicate the predominance of inverse γ -turns.

As for cyclic pohlianin A in 100% ACN-*d*₃, three conformers (**1.1**–**1.3**, Fig. 5) were found to populate the sample according to the energy criterion used (>1% Boltzmann population, Table S3†). Calculations were performed at different levels, including B3PW91/PCM(ACN)/6-31G(d,p), B3PW91/PCM(ACN)/cc-pVDZ B3LYP/PCM(ACN)/6-31G(2d,p), M06-2X/PCM(ACN)/6-31G(d,p), B3LYP/PCM(ACN)/cc-pVDZ, and wB97XD/PCM(ACN)/6-31G(d), without any isotopic exchange.

It is not unusual to macrocycles to have few predominant conformers despite the apparent flexibility.^{30,31} The excellent correlation between experiment and calculations at the B3LYP/PCM(ACN)/cc-pVDZ level (Fig. 6) confirms the conformational ensemble found. Interestingly, the second lowest-energy conformer identified (**1.2**) presented almost identical dihedral angles and secondary structure motifs to those described in DMSO following NMR analysis⁷ (Table S1†), that is, types I and VIa β -turns with a β -bulge feature. This is an important validation of the conformational search used, since, as described in the calculations section, MD runs were started at different geometries and did not consider the main conformation derived from NMR. While the original NMR analysis resulted in a single averaged conformation, the VCD investigation led to identification of two further conformations that are significantly populated in both ACN-*d*₃ and 75 : 25 ACN-*d*₃/D₂O solutions. These results are pivotal to understanding structure–activity relationships in orbitides. VCD revealed significant contributions of conformer **1.1** that presents the same type I β -turn at Val₄–Leu₅ as **1.2**, however, with 2 additional classic γ -turns involving Gly₃ and Leu₆ in the $i + 1$ position. The third identified conformer (**1.3**), on the other hand, maintains the type VIa1 β -turn at Tyr₇–Pro₁ and the β -bulge feature present in

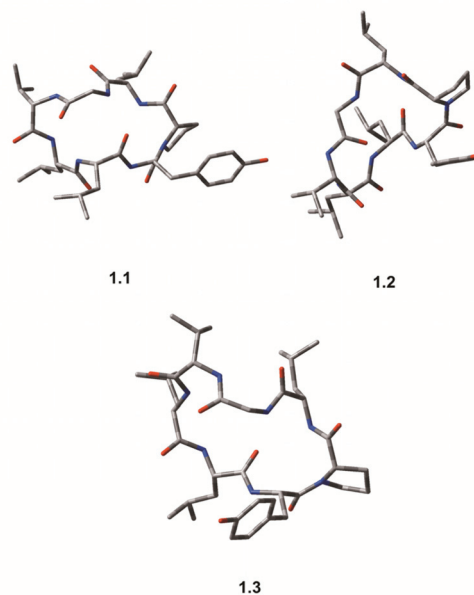


Fig. 5 Lowest-energy conformers of **1** (cyclic) identified at the B3PW91/PCM(ACN)/cc-pVDZ level and used as a single average. Hydrogen atoms omitted for clarity.

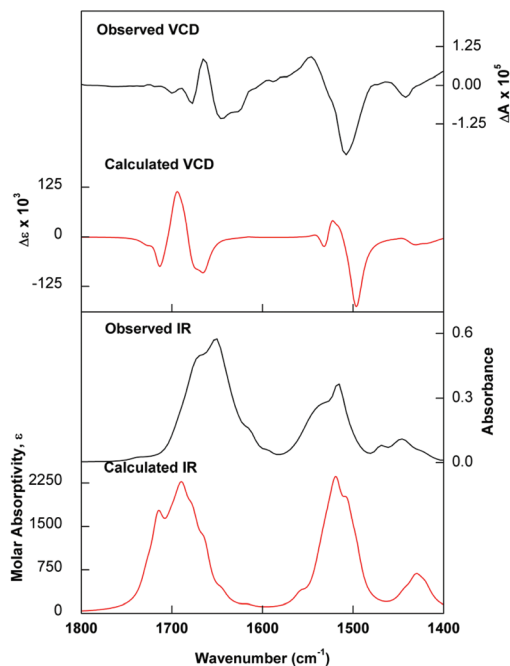


Fig. 6 Comparison of experimental IR and VCD spectra of **1** (cyclic) recorded in 100% ACN- d_3 (black trace) with calculated [B3PW91/PCM(ACN)/cc-pVDZ] IR and VCD spectra for the simple average of the three lowest-energy conformers identified for all- L **1** (cyclic) (red trace).

1.2 but a change in the glycine dihedral angles precludes the formation of the type I β -turn at Val₄–Leu₅ and leads to a classic γ -turn instead (Gly₃ in the $i + 1$ position). Additionally, conformer **1.3** does not present the tyrosyl aromatic ring over the proline ring, as observed in **1.1** and **1.2**. The prevalent stacking of the pyrrolidine ring of proline and the aromatic ring of tyrosine is in accordance with NMR results (α and β protons of proline significantly shifted upfield). It is noteworthy that even though the individual conformers **1.1** and **1.2** would present acceptable agreements with experiment (Fig. S7†), the correct relative intensities and signs of both amide I and II are only obtained taking into account the three conformers identified. The amide I VCD spectral features of pohlianin A in ACN- d_3 arise mainly from out-of-phase C=O stretching of Leu₅, Leu₆, and Tyr₇ (negative band at around 1640 cm⁻¹); out-of-phase and in-phase C=O stretching of Leu₂, Gly₃, Leu₆ and Tyr₇ (positive band at around 1660 cm⁻¹); and out-of-phase C=O stretching of Pro₁, Leu₂, Gly₃, and Val₄ (negative band at around 1680 cm⁻¹). The amide II negative VCD band at 1510 cm⁻¹ involves mainly the type I β -turn residues Val₄, Leu₅, Leu₆. The positive amide II band at 1545 cm⁻¹ arises mainly from vibrations of Leu₂, Gly₃, and Tyr₇. These amide II results are in accordance with the VCD signature of type I β -turns (+,– couplet) established for homochiral di-²⁵ and tetrapeptides,²⁴ despite the presence of additional γ - and type VIa1 β -turns in **1**. The mixed occurrence of β -turns and classic γ -turns, however, gives rise to the –,+,– amide I feature in the cyclic heptapeptide. The negative VCD band at around

1440 cm⁻¹ is related to O=C–N and C _{α} –C=O stretchings of Tyr₇ bonded to Pro₁, as well as CH₂ scissoring and C _{α} –H bendings of both Pro₁ and Tyr₇ and aromatic deformation modes of Tyr₇. This negative feature may be considered as a marker for the stacking of the pyrrolidine ring of proline and the aromatic ring of tyrosine, since conformer **1.3** presents a positive band in this region (Fig. S8†). The additional spectral simulations at the B3PW91/PCM(ACN)/6-31G(d,p), B3LYP/PCM(ACN)/6-31G(2d,p), M06-2X/PCM(ACN)/6-31G(d,p), B3LYP/PCM(ACN)/cc-pVDZ, and wB97XD/PCM(ACN)/6-31G(d) levels (Fig. S6†) resulted in similar IR/VCD amide I and II profiles for the orbitide. The main discrepancies were observed using the M06-2X and wB97XD hybrid functionals, especially in the amide I VCD bands. On the other hand, the overall IR agreement with experiment, as well as that of the amide II VCD band, was improved with wB97XD. As calculations of systems of this size are computationally expensive, one may choose any of the levels presented herein based on the computational resources available.

Conclusions

The biosynthesis of orbitides seems to take place *via* ribosomally produced long chain linear precursors that are subjected to proteolytic cleavage steps followed by cyclization.⁵ Even though the conformational analyses performed herein for both linear and cyclic pohlianin A cannot be directly used to draw conclusions on the orbitide biosynthesis, knowing the preferred structures of both forms in partially aqueous solution may improve the understanding of their bioactivity requirements. IR and VCD investigations aided by MD and DFT calculations led to the identification of inverse γ -turns as the most prevalent structural motif in the linear precursors, while type I and type VIa β -turns with a β -bulge are commonly found in the cyclic peptide, along with some classic γ -turns. Although the γ -turn in **1a.2** involving Leu₆ in the $i + 1$ position seems to be conserved during the cyclization process, it changes from inverse to a classic type following head-to-tail linkage. The VOA results obtained for pohlianin A were in agreement with the NMR-derived conformational predictions for a series of cyclic heptapeptides, however, it revealed significant populations of two additional conformers with distinct secondary structures. Since bioactivity is highly dependent on the 3D structures adopted by a given compound in solution, VCD is presented as powerful and validated tool that provides conformational and configurational information to be used in structure–activity relationship studies of orbitides. As turns constitute the third most common secondary structural elements in proteins, systematic VCD studies of turn-rich orbitides, such as the one presented herein, contribute to a more precise prediction and characterization of both β - and γ -turns. Finally, the DFT calculations of the heptapeptide pohlianin A in different levels represent a first effort to benchmark functional and basis sets for IR/VCD predictions of these large cyclic peptides.

Experimental

Peptide synthesis

The synthetic linear precursors of pohlianin A (VLLYPLG) in both configurations *D* and *L* were manually synthesized according to the standard $N\alpha$ -Fmoc protecting group strategy.³² The side chain protecting group *t*Bu (*t*-Butyl) was used only for the Fmoc-amino acid Tyr, while the other Fmoc-amino acids were used unprotected. The first step was the α -amino group deprotection of the Rink-amide-MBHAR, performed in 20% 4-methylpiperidine/dimethylformamide (DMF) for 1 min and then 20 min. After this, the coupling of the Fmoc-amino acid by its α -carboxyl group was evaluated at threefold excess, using *N,N'*-diisopropylcarbodiimide (DIC)/1-hydroxybenzotriazole (HOBT) in 50% (v/v) DCM (methylene chloride)/DMF or, when a Fmoc-amino acid recoupling was needed, *N,N,N',N'*-tetramethyl-*O*-(1*H*-benzotriazol-1-yl) uronium hexafluorophosphate (HBTU)/diisopropylethylamine (DIEA) in 50% (v/v) DCM/*N*-methylpyrrolidone (NMP) was employed. After 2 hours of coupling, the ninhydrin test was performed to monitor the completeness of the reaction. For removal of the peptides from the matrix support and the side protecting groups, the cleavage step was simultaneously performed with 95% trifluoroacetic acid (TFA), 2.5% triisopropylsilane (TIS), and 2.5% Milli-Q water for 2 hours, under stirring. After this procedure, the crude peptides were precipitated with anhydrous ethyl ether, separated from soluble non-peptide material by centrifugation, and extracted into a 30% acetonitrile/H₂O solution (v/v). After this step, the peptides were lyophilized. Purification was performed by semi-preparative HPLC, model Prominence (Shimadzu, Tokyo, Japan), with a Shim-Pack reverse phase C-18 column, model PREP-ODS, dimensions of 20 × 250 mm and particle size of 15 μ m, using solvents A (0.045% TFA:H₂O) and B (0.036% TFA:ACN) with a linear gradient of 5–95% (v/v) of solvent B in A for 30 min, with a flow rate of 5 mL min⁻¹. Peptide homogeneity and purity were checked by analytical HPLC using a Shim-pack C-18 column, model VP-ODS, dimensions of 4.6 × 250 mm and particle size of 5 μ m, at oven temperature of 35 °C, and with the same linear gradient program, with a flow rate of 1.0 mL min⁻¹ and UV detection at 220 nm. The cyclizations were performed by diluting the linear peptides at a concentration of 1 mmol L⁻¹ in DMF, where a usual coupling reaction was carried out with 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxide hexafluorophosphate (HATU)/HOBT/DIEA (1.5; 1.5; 4.5 equivalents, per 2 h). In this reaction, the amino group of valine at position 1 acted as a nucleophile to displace the hydroxyl group of glycine at position 7, forming a peptide bond (head-to-tail cyclization). The reaction was carried out under conditions of high dilution (1 mmol L⁻¹, peptide in DMF) in order to minimize unwanted intermolecular reactions, such as oligo and polymerizations, in addition to improving the yield of cyclic peptide formation. The presence of the linear and cyclic peptides was confirmed by mass spectrometry using a Bruker Amazon SL Ion trap Electrospray Mass Spectrometer (Billerica, MA, USA). All peptides used presented a high degree of purity (>95%).

VCD spectroscopy

IR and VCD spectra of *L*- and *D*-**1** and *L*- and *D*-**1a** were recorded with a single-PEM ChiralIR-2X FT-VCD spectrometer (BioTools, Inc., Jupiter, FL, USA) using a resolution of 8 cm⁻¹ and a collection time of 20 h. The optimum retardation of the ZnSe photoelastic modulator (PEM) was set at 1600 cm⁻¹. The IR and VCD spectra were recorded in solution using 75 : 25 ACN-*d*₃/D₂O and 100% ACN-*d*₃ for **1** and only 75 : 25 ACN-*d*₃/D₂O for **1a**. Concentrations were of 75 mg mL⁻¹ for **1** and 100 mg mL⁻¹ for **1a** and the measurements were carried out in a BaF₂ cell with a 50 μ m path length. Minor instrumental baseline offsets were eliminated from the final VCD spectrum of **1** and **1a** by taking the half difference of the VCD spectra of both enantiomers obtained under identical conditions.

Calculations

Molecular dynamics (MD) simulations were performed for all-**1** (cyclic) and **1a** (linear) in periodic box conditions with approximately 1000 explicit water molecules for **1** and 700 water molecules for **1a** using the AMBER force field as implemented in Hyperchem 8.0.10 software package. A single molecule of either **1** or **1a** was placed at the centre of the box. Different starting geometries were used as input for multiple MD simulations using the following parameters: total run time of 50 μ s, time step of 0.001 ps, simulation temperature of 300 K. A combination of short MD runs was used (*i.e.* 500 ps) and their final structures were used as inputs for the following simulations. The different conformers obtained from individual snapshots taken at approximately every 10 ps were geometry optimized and had harmonic frequencies calculated at DFT level. All DFT calculations were carried out at 298 K in water, DMSO, or ACN solutions using the polarizable continuum model (PCM) in its integral equation formalism version (IEFPCM) incorporated in Gaussian 09 software.³³ Geometry optimizations were performed initially at the B3PW91/PCM/6-31G(d,p). The conformers contributing with $\geq 1\%$ of the Boltzmann distribution, considering Gibbs free energies, were selected for IR/VCD spectral calculations. IR and VCD spectra **1** and **1a** were obtained using dipole and rotational strengths from Gaussian software, which were calculated at the B3PW91/PCM(ACN)/cc-pVDZ, B3LYP/PCM(ACN)/6-31G(2d,p), M06-2X/PCM(ACN)/6-31G(d,p), B3LYP/PCM(ACN)/cc-pVDZ, B3PW91/PCM(ACN)/6-31G(d,p), and wB97XD/PCM(ACN)/6-31G(d) levels for **1**, and at the B3PW91/PCM(DMSO)/6-31G(d,p) and B3PW91/PCM(DMSO)/cc-pVDZ levels for **1a**, and converted into molar absorptivities (M⁻¹ cm⁻¹). The partial replacement of hydrogen atoms with deuterium atoms in **1a**, as well as the creation of zwitterionic forms, were carried out using GaussView 5.0.9 software. For **1a**, DFT calculations were also performed using the two-layer ONIOM method. The high layer/QM region was treated at the B3LYP/6-31G(d) level of theory for geometry optimizations and harmonic frequency calculations. The low layer/MM water region was treated with AMBER force field without electronic embedding. Each spectrum, calculated using the different approaches described

above, was plotted as a sum of Lorentzian bands with half-widths at half-maximum of 8 cm^{-1} for **1** and 10 cm^{-1} for **1a**. The calculated wavenumbers were multiplied with a scaling factor of 0.98 and the single averaged composite spectra were plotted using Origin software. For **1** and **1a**, the spectra obtained in different levels of theory can be found in the ESI.†

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by grants from the São Paulo Research Foundation – FAPESP (grants# 2016/00446-7 and 2019/22319-5) and the National Council for Scientific and Technological Development – CNPq (grant# 431978/2018-2). MASY thanks CNPq for the research scholarship. HRLS, LEVNB and EFV thank the research group “Peptides: Synthesis, Optimization and Applied Studies – PeSEAp” and Laboratório de Equipamentos Multiusuários (LEMU) from the School of Sciences and Engineering (Unesp). This research was also supported by resources supplied by the Centre for Scientific Computing (NCC/GridUNESP) of São Paulo State University (UNESP).

References

- P. G. Arnison, M. J. Bibb, G. Bierbaum, A. A. Bowers, T. S. Bugni, G. Bulaj, J. A. Camarero, D. J. Campopiano, G. L. Challis, J. Clardy, P. D. Cotter, D. J. Craik, M. Dawson, E. Dittmann, S. Donadio, P. C. Dorrestein, K.-D. Entian, M. A. Fischbach, J. S. Garavelli, U. Göransson, C. W. Gruber, D. H. Haft, T. K. Hemscheidt, C. Hertweck, C. Hill, A. R. Horswill, M. Jaspars, W. L. Kelly, J. P. Klinman, O. P. Kuipers, A. J. Link, W. Liu, M. A. Marahiel, D. A. Mitchell, G. N. Moll, B. S. Moore, R. Müller, S. K. Nair, I. F. Nes, G. E. Norris, B. M. Olivera, H. Onaka, M. L. Patchett, J. Piel, M. J. T. Reaney, S. Rebuffat, R. P. Ross, H.-G. Sahl, E. W. Schmidt, M. E. Selsted, K. Severinov, B. Shen, K. Sivonen, L. Smith, T. Stein, R. D. Süßmuth, J. R. Tagg, G.-L. Tang, A. W. Truman, J. C. Vederas, C. T. Walsh, J. D. Walton, S. C. Wenzel, J. M. Willey and W. A. van der Donk, *Nat. Prod. Rep.*, 2013, **35**, 108.
- M. F. Fisher, C. D. Payne, J. Rosengren and J. S. Mylne, *J. Nat. Prod.*, 2019, **82**, 2152.
- S. D. Ramalho, M. E. F. Pinto, D. Ferreira and V. S. Bolzani, *Planta Med.*, 2018, **84**, 558.
- R. Jwad, D. Weissberger and L. Hunter, *Chem. Rev.*, 2020, **120**, 9743.
- J. R. Chekan, P. Estrada, P. S. Covello and S. K. Nair, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 6551.
- P. M. Cândido-Bacani, P. O. Figueiredo, M. F. C. Matos, F. R. Garcez and W. S. Garcez, *J. Nat. Prod.*, 2015, **78**, 2754.
- C. Auvin-Guette, C. Baraguey, A. Blond, H. S. Xavier, J.-L. Pousset and B. Bodo, *Tetrahedron*, 1999, **55**, 11495.
- K. Thell, T. Hellinger, G. Schabbauer and C. W. Gruber, *Drug Discovery Today*, 2014, **19**, 645.
- S. C. Barbosa, T. M. Nobre, D. Volpati, P. Ciancaglini, E. M. Cilli, E. N. Lorenzón and O. N. Oliveira Jr., *Colloids Surf., B*, 2016, **148**, 453.
- J. Tian, Y. Shen, X. Yang, S. Liang, L. Shan, H. Li, R. Liu and W. Zhang, *J. Nat. Prod.*, 2010, **73**, 1987.
- M. E. F. Pinto, J. M. Batista Jr., J. Koehbach, P. Gaur, A. Sharma, M. Nakabashi, E. M. Cilli, G. M. Giesel, H. Verli, C. W. Gruber, E. W. Blanch, J. F. Tavares, M. S. Silva, C. R. S. Garcia and V. S. Bolzani, *J. Nat. Prod.*, 2015, **78**, 374.
- S. D. Ramalho, M. E. F. Pinto, R. K. Andricopulo, P. R. S. Sanches, E. R. Silveira, E. M. Cilli, A. D. Andricopulo and V. S. Bolzani, *J. Braz. Chem. Soc.*, 2019, **30**, 2153.
- J. Damjanovic, J. Miao, H. Huang and Y.-S. Lin, *Chem. Rev.*, 2021, **121**, 2292.
- S. D. Ramalho, C. K. Wang, G. J. King, K. A. Byriel, Y.-H. Huang, V. S. Bolzani and D. J. Craik, *J. Nat. Prod.*, 2018, **71**, 2436.
- T. A. Keiderling, *Chem. Rev.*, 2020, **120**, 3381.
- H. R. Wyssbrod and M. Diem, *Biopolymers*, 1992, **32**, 1237.
- P. Xie and M. Diem, *J. Am. Chem. Soc.*, 1995, **117**, 429.
- P. Xie, W. Zhou and M. Diem, *J. Am. Chem. Soc.*, 1995, **117**, 9502.
- J. Hilario, J. Kubelba, F. A. Syud, S. H. Gellman and T. A. Kiederling, *Biopolymers*, 2002, **67**, 233.
- P. Bouř, J. Kim, J. Kaptán, R. P. Hammer, R. Huang, L. Wu and T. A. Kiederling, *Chirality*, 2008, **20**, 1104.
- M. P. D. Hatfield, R. F. Murphy and S. Lovas, *Biopolymers*, 2010, **93**, 442.
- E. Vass, Z. Majer, K. Köhalmly and M. Hollósi, *Chirality*, 2010, **22**, 762.
- C. Merten, F. Li, K. Bravo-Rodriguez, E. Sanchez-Garcia, Y. Xu and W. Sander, *Phys. Chem. Chem. Phys.*, 2014, **16**, 5627.
- N. Berger, F. Li, B. Mallick, J. T. Brüggemann, W. Sander and C. Merten, *Biopolymers*, 2017, **107**, 28.
- T. Vermeyen and C. Merten, *Phys. Chem. Chem. Phys.*, 2020, **22**, 15640.
- M. Migliore, A. Bonvicini, V. Tognetti, L. Guilhaudis, M. Baaden, H. Oulyadi, L. Joubert and I. Ségalas-Milazzo, *Phys. Chem. Chem. Phys.*, 2020, **22**, 1611.
- E. F. Vicente, L. G. M. Basso, G. F. Cespedes, E. N. Lorenzón, M. S. Castro, M. J. S. Mendes-Giannini, A. J. Costa-Filho and E. M. Cilli, *PLoS One*, 2013, **8**, e60818.
- J. Kessler, V. Andrushchenko, J. Kaptán and P. Bouř, *Phys. Chem. Chem. Phys.*, 2018, **20**, 4926.
- L. G. Gagliardi, C. B. Castells, C. Ràfols, M. Rosés and E. Bosch, *J. Chem. Eng. Data*, 2007, **52**, 1103.

- 30 W. F. Altei, D. G. Picchi, B. M. Abissi, G. M. Giesel, O. Flausino Jr., M. Reboud-Ravaux, H. Verli, E. Crusca Jr., E. R. Silveira, E. M. Cilli and V. S. Bolzani, *Phytochemistry*, 2014, **107**, 91.
- 31 D. P. Demarque and C. Merten, *Chem. Commun.*, 2020, **56**, 10926.
- 32 L. A. Carpino and G. Y. Han, *J. Org. Chem.*, 1972, **37**, 3404.
- 33 M. J. Frisch, G. Trucks, H. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. Hratchian, A. Izmaylov, J. Bloino, G. Zheng, J. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. Montgomery, J. Peralta, F. Ogliaro, M. Bearpark, J. Heyd, E. Brothers, K. Kudin, V. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. Burant, S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. Millam, M. Klene, J. Knox, J. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. Stratmann, O. Yazyev, A. Austin, R. Cammi, C. Pomelli, J. Ochterski, R. Martin, K. Morokuma, V. Zakrzewski, G. Voth, P. Salvador, J. Dannenberg, S. Dapprich, A. Daniels, Ö. Farkas, J. Foresman, J. Ortiz, J. Cioslowski and D. Fox, *Gaussian 09, Revision A.01*, Gaussian, Inc., Wallingford CT, 2009.