THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Conformational Preference of Fused Carbohydrate-Templated Proline Analogues—A Computational Study

Robel B. Teklebrhan,[†] Neil W. Owens,[†,∥] James D. Xidos,[†] Georg Schreckenbach,*[,†] Stacey D. Wetmore,[‡] and Frank Schweizer*[,†,§]

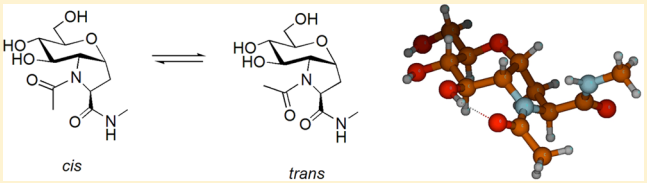
[†]Department of Chemistry, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada

[‡]Department of Chemistry & Biochemistry, University of Lethbridge, Lethbridge, Alberta T1K 3M4, Canada

[§]Department of Medical Microbiology, University of Manitoba, Winnipeg, Manitoba R3A 1R9, Canada

Supporting Information

ABSTRACT: The structure of fused *C*-glucosylproline hybrid (*GlcProH*) has been studied in detail computationally. A systematic molecular mechanics/Monte Carlo search has been performed in order to cover the entire conformational space of *GlcProH*. This has been followed by density-functional (DFT B3LYP) calculations in the gas phase and in aqueous solution, using the polarizable continuum model (PCM). In the gas phase, a large excess of the *cis* conformation with respect to the prolyl amide bond is found. This is reversed in aqueous solution where the calculations show 80% *trans* conformers, which is in accordance with experimental data. Thus, the PCM model is capable of accurately predicting *cis*–*trans* ratios. The free energy of solvation is not correlated with the dipole moment. Hence, a model (such as PCM) is required that takes into account the complete charge distribution. The reversal of the *cis*–*trans* ratio between gas phase and solution also emphasizes the effects of different free energies of solvation for the distinct conformers. Nevertheless, the energy difference between the *cis* and *trans* conformers is very small in solution (0.18 kcal/mol). Intramolecular hydrogen bonding is found to stabilize the *cis* conformers exclusively, which is a result of the rigid geometry of the fused rings. This can be contrasted to related more flexible molecules that show hydrogen bonding for both *cis* and *trans* isomers. The hydrogen bonding is at least partially responsible for the preferential stabilization of the *cis* conformers in the gas phase and a very small *cis*–*trans* energy difference in solution.



INTRODUCTION

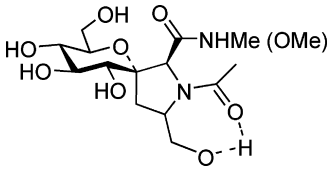
Among the 20 essential amino acids, proline (Pro) is unique since it possesses a cyclic side chain that is fused to the amino function, forming a five-membered ring. The five-membered ring severely reduces the conformational freedom of proline by limiting rotations around the ϕ dihedral angle. This, in turn, reduces the energy difference between the prolyl amide *cis* and *trans* isomers, to the point that they have very similar energies. Thus, contrary to all the other amino acids that exist almost exclusively in their *trans* conformer, proline can also exist in the *cis* form. For this reason, proline is crucial, for instance, for inducing a reversal in the peptide backbone conformation.[1,2] Proline *cis*–*trans* isomerization is also relevant since it constitutes the rate-determining step in the folding pathway of various proteins and peptides.[3]

In order to better understand—and possibly modulate—the behavior of proteins and peptides, it is desirable to identify ways to modify the prolyl *cis*–*trans* isomer ratio using proline derivatives. With this goal in mind, a number of proline derivatives have been proposed over the years, such as silaproline,[4] azaprolines,[5,6] C^{β}-, C^{γ}-, and C^{δ}-substituted prolines,[7−17] pseudoprolines,[18−22] proline analogues with different ring size,[18,23,24] or fused bicyclic proline.[25−33]

Despite the variety of proline analogues synthesized so far, none are capable of shifting the prolyl amide equilibrium toward either the *cis* or *trans* isomer by making only a simple chemical substitution. In other words, completely different proline analogues would be required to achieve either a *trans* or a *cis* bias. With this limitation in mind, we have embarked on developing simple proline analogue building blocks with strategically located functional groups in which the prolyl *cis*–*trans* isomerization could be tuned through relatively simple substitutent chemistry, even after incorporation into a full peptide.[33] Recently, we have reported the synthesis of two spirocyclic glucosyl-(3(*S*)-hydroxy-5(*S*)-hydroxymethyl)proline and glucosyl-(3(*S*)-hydroxy-5(*R*)-hydroxymethyl)proline hybrids, Scheme 1.[31] Incorporation of both proline building blocks into model peptides via acylation of the N-terminal amino group allowed us to study the conformational properties of these proline analogues.[30] The thermodynamics and kinetics of the prolyl amide *cis*–*trans* isomerization in these peptide mimics were investigated in detail with a large variety of techniques,[30] including an extensive computational study.[30,32]

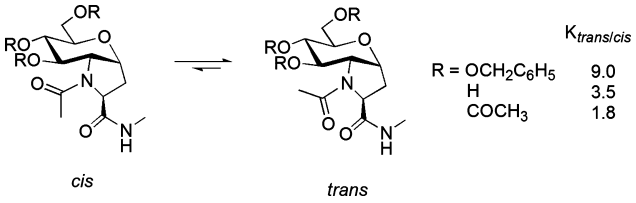
Received: October 29, 2012
Revised: December 9, 2012
Published: December 10, 2012

Scheme 1. Spirocyclic Glucose-Proline Hybrid



In a different, though closely related study, we have reported on the synthesis and characterization of a fused bicyclic *C*-glucosylproline hybrid *N-acetyl-GlcProH-NHMe* (*GlcProH*) that was incorporated into peptide mimics such as *AcN-GlcProH-NHMe*, Scheme 2.[33] Significant changes in the *cis*−*trans* isomer

Scheme 2. Substituent Effects of *C*-Glucosylproline Hybrid *GlcProH* on Prolyl Amide *trans/cis* ratio ($K_{trans/cis}$) Isomerization in Model Peptide *AcN-GlcProH-NHMe* (R = H in the current study)



equilibrium could be achieved through modification of the carbohydrate scaffold; however, the exact mechanism by which this is achieved was unclear. In the current paper, we present the results of a detailed computational study of this compound. Bicyclic *GlcProH* rigidly combines the molecular elements of carbohydrates (pyran-based polyol) with the unique features of proline. The resulting *GlcProH* is a rigid, polyfunctional building block that may find use as a proline mimetic, a glycomimetic, or a scaffold for combinatorial synthesis.

Several conformationally constrained L-proline analogues have been developed,[34,35] but recently, fused bicyclic prolines have attracted interest due to their increased rigidity, which may permit better control of the prolyl amide *cis*−*trans* ratio.[27−29] For instance, bicyclic proline analogues have been incorporated into bioactive peptides including angiotensin[36] and thyroliberin.[37] Because of its stable chair conformation, glucose provides an ideal scaffold to template proline, since it freezes the orientation of four proline atoms (C_β, C_γ, C_δ, and N). Moreover, the polyhydroxylated sugar scaffold in *GlcProH* can participate in intramolecular hydrogen bonding and electrostatic interactions, which can induce novel or unusual secondary structures in *GlcProH*-containing peptides. For instance, incorporation of polyhydroxylated sugar amino acids into small peptides, such as opioid peptides[38] and gramicidin S,[39] prohibited the formation of the targeted secondary structural motif. Appearing instead were unusual turn structures stabilized by intramolecular hydrogen bonds between sugar hydroxyl groups and peptidic amide backbone.[40] In this context, it is interesting to note that hydroxylation of proline and lysine is a crucial posttranslational modification that controls the stability of the triple helix in structural proteins such as collagens. Furthermore, derivatization of the sugar hydroxyl groups in *GlcProH* provides a unique tool to control prolyl amide *cis*−*trans* isomerization in peptides (Scheme 2), which may find future applications in the design of artificial collagens and biomaterials.

The purpose of the current paper is to provide the first detailed, computational structural study of *GlcProH*. A specific computational protocol for achieving these goals, which was previously used to study model peptides including the spirocyclic compound *Ac-Glc3(S)-HypH-OMe*,[30] is applied. Our approach uses accurate density functional (DFT) methods and continuum solvation models and ensures comprehensive coverage of the large conformational space. Computations in the gas phase and solution are compared, and the role of intramolecular hydrogen bonding is critically evaluated, since this has been shown to dictate the structural features of other proline analogues.

COMPUTATIONAL DETAILS

Over the years, proline and its derivatives have been the subject of a relatively large number of computational studies at various levels of theory.[5−7,13,17,21,24,30,32,41−64] Computational studies of such flexible organic molecules are challenging due to the large conformational space that grows exponentially with the degrees of freedom (and therefore number of atoms).[65] A proper computational method must ensure that this conformational space is adequately sampled, since the neglect of just one low-energy conformer might make the results completely unreliable. Moreover, an appropriate level of theory is required to capture the effects of the electronic structure, and accurate methodologies must account for the presence of a nonpolar or polar solvent.[66] In our previous paper, we addressed these challenges and developed a computational protocol for studying flexible organic molecules such as *GlcProH*.[32] The same computational protocol has been applied in the current study and is described in detail below.

Initially, molecular mechanics (MM) searches were carried out using the Spartan '02 program package.[67] Specifically, we used the Monte Carlo[68,69] (MC) random searching algorithm and systematic search methods[65,68] with the MMFF94 force field[70−76] as implemented in Spartan, where MC was applied to cyclic systems. A combination of MC and systematic search was applied in a "Build and Search" approach. In this approach, the search starts with a small molecule (that constitutes a fraction of the target molecule) with few degrees of freedom, which in the present case was the five-membered ring (CH_2−CH_2−CH_2−CH_2−NH). Its conformational space is mapped out using systematic search and MC methods. The search is performed five times using MC and five times using systematic search to account for its completeness, i.e., until a redundant outcome (no more new conformers) is obtained. This results in a number of unique conformers for this initial small molecule. These conformers are used as starting structures for building a sequentially larger molecule, which in the present case was obtained by attaching the six-membered ring CH_2−CH_2−CH_2−CH_2−CH_2−O to each conformer. Systematic search and MC methods are again applied in a similar manner to obtain all possible conformers of this larger molecule. Finally, the target compund was built onto each of the conformers that resulted from these searches. Redundant, duplicate conformers of the target molecule, detected by superimposing all of the conformers, are discarded. In this way, a total of 1200 possible structures were found; of these, 101 unique structures were determined at the MMFF94 level of theory. The overall procedure ensures a consistent and redundant data set with a high confidence level.

These 101 unique conformers were used as input structures for density functional theory[77] (DFT) optimizations. All DFT

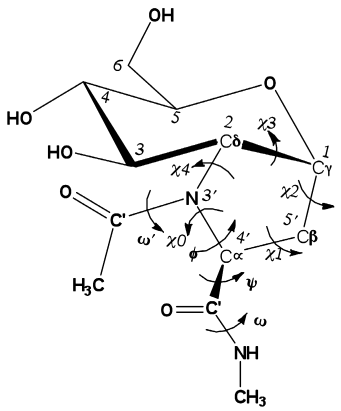
dx.doi.org/10.1021/jp310690c | J. Phys. Chem. B 2013, 117, 199−205

calculations were performed with the Gaussian03 (g03) program.[78] We applied the B3LYP[79−81] level of theory with 6-31+G(d,p) basis sets. Full geometry optimizations were performed in the gas phase, followed by frequency calculations. In this manner, all optimized stationary points were characterized as local minima on the potential energy surface (PES). The optimized gas-phase structures were reoptimized in water using the PCM continuum solvation model.[82] Default radii were used as implemented in g03. Frequency calculations in PCM were also performed, and all stationary points were again characterized as local minima on the PES. Finally, Boltzmann statistics at 298.15 K (25.0 °C) were applied to the gas phase and solution conformers in order to obtain the respective conformer distributions and the prolyl amide *cis*/*trans* ratios.

Boltzmann statistics should in principle be based on relative free energies, based on the equation $G_{soln} = E_{soln} + G_{nes} + \Delta G_{corr\_gas}$, where E_{soln} is the electronic energy of the solute in the presence of the continuum solvent field, G_{nes} is the sum of nonelectrostatic contributions to the solvation free energy, and ΔG_{corr_gas} is the thermal contributions evaluated in the gas phase. Given the approximations made in the evaluation of this term and the relatively low level of theory applied, following our earlier work,[32] we have chosen not to include ΔG_{corr_gas}, since the error in including this term may be more significant than leaving it out. Moreover, as is shown by Flores-Ortega et al.[83] in their study of closely related systems, the inclusion of this term has a relatively small effect on the final results. This is expected, since the interconversion between conformations is an isodesmic process.

The calculated gas phase and solution conformers were analyzed in order to detect trends in the geometries and electronic structure. Specifically, as part of the analysis, we have characterized the backbone torsion and endocyclic angles. The definitions of the various angles are provided in Scheme 3. The definitions are completely analogous to those used in our previous work.[32] Furthermore, following the procedure applied in our previous work,[32] we investigated the most stable conformers for internal hydrogen bonding, using a bond length cutoff of ≤2.5 Å as a criteria for the existence of a hydrogen bond.[84] (This refers to the distance between the hydrogen atom and an adjacent heavy atom.) The nature of the different conformers was characterized (classified) on the basis of the backbone torsion angles ϕ and ψ.[55,85−87] The details of the nomenclature are provided in the Supporting Information.[32]

Scheme 3. Definitions of the Backbone and Endocyclic Torsion Angles



RESULTS AND DISCUSSION

Conformational Distribution of *N-acetyl-GlcProH-NHMe* (*GlcProH*) in the Gas Phase and in Water. Of the 101 unique conformers determined, 15 conformers in the gas phase and 14 conformers in solution each contribute at least 0.5% to the overall distribution, according to the Boltzmann statistics. Details for these are given in Tables S1 and S2 of the Supporting Information. Taken together, these conformers make up 99.99% (gas phase) and 99.84% (water), respectively, of the total distributions. In other words, the contributions from all of the higher energy conformers are essentially negligible. Using the classification of the backbone and endocyclic torsion angles and the prolyl *cis*−*trans* isomerization, eight types of local minima were obtained in the gas phase. Specifically, these are c-α_L[d], c-ε_L[d], t-γ_L[d], c-ε_L[u], t-γ_L[u], c-ε_L[u]*, c-γ_L[d], and c-g_L[u]. However, only two types of local minima, t-ε_L[d] and c-ε_L[d], were found in water, representing the prolyl amide *trans* and *cis* conformations. This change from gas phase to solution clearly illustrates the need for full optimizations in the solvent environment. Representative examples of the optimized structures are provided in Figures 1 and 2, respectively. The relative energies of the conformations in the gas phase and water follow the order of c-α_L[d], c-ε_L[d] < t-γ_L[d] < c-ε_L[u] < t-γ_L[u] < c-ε_L[u]* < c-γ_L[d] < c-g_L[u], and t-ε_L[d] < c-ε_L[d], respectively, with the lowest relative energy belonging to the most stable conformer. Thus, the most stable *trans* and *cis* isomers in the gas phase are c-α_L[d] and t-γ_L[d], respectively. The energy difference between the most stable *cis* and *trans* conformers in the gas phase was calculated to be 1.53 kcal/mol, which is far larger than that for the aqueous solution (0.18 kcal/mol).

Prolyl Amide *cis*−*trans* Distributions. The calculated conformational *cis*−*trans* distributions are provided in Table 1 where they are also compared to the experimental values.[33] More detailed information for the most important conformers (those that contribute at least 0.5% to the overall conformer distribution) is provided in Tables S1 and S2 of the Supporting Information for the gas phase and aqueous phase calculations, respectively.

The prolyl amide conformational distributions calculated for the gas phase and water (aqueous solution) have opposite trends; see Table 1. Indeed, in the gas phase, the dominating conformation is *cis* (95%), whereas, in water, it is *trans* (80%). The calculated *cis* percentage in water (20%) is in good agreement with the experimentally observed value of 13%. This kind of agreement was also found for the spirocylic carbohydrate templated proline derivatives in our previous study.[32] We believe this agreement between solvent phase calculations and experiment should be considered close, given the simplified treatment of the—in principle rather complex—solvent structure in the continuum solvation model.[88−90] We also note the absence of explicit solute−solvent hydrogen bonds in the continuum solvation models. Overall, comparison of the gas phase and solution results for the conformer distribution demonstrates the crucial importance of including solvation in the modeling.

Geometrical Parameters. Key geometrical parameters (bond lengths and angles) for the two most stable *cis* and *trans* isomers of *GlcProH* in water are provided in Tables 2 (angles) and S3 (Supporting Information, bond lengths). The

dx.doi.org/10.1021/jp310690c | J. Phys. Chem. B 2013, 117, 199−205

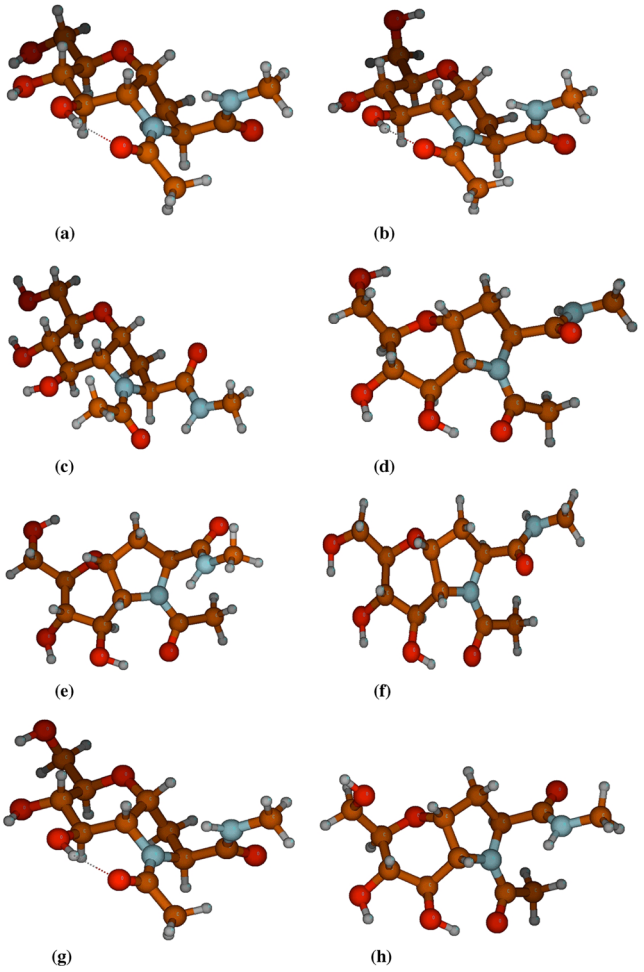


Figure 1. Representative Gas-Phase Optimized Structures of the Most Stable Conformers of *N-acetyl-GlcProH-NHMe (GlcProH)*: (a) c-α_L[d]; (b) c-ε_L[d]; (c) t-γ_L[d]; (d) c-ε_L[u]; (e) t-γ_L[u]; (f) c-ε_L[u]*; (g) c-γ_L[d]; (h) c-γ_L[u].

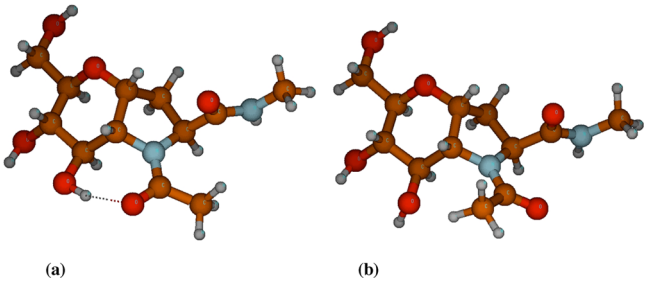


Figure 2. Representative solution-phase (water) optimized structures of the most stable conformers of *N-acetyl-GlcProH-NHMe (GlcProH)*: (a) c-ε_L[d]; (b) t-ε_L[d].

Table 1. Calculated[a] and Experimental Total Population Distributions (%)

	gas phase	in H_2O	exp., in H_2O[b]
cis isomers	94.75	19.95	13
trans isomers	5.25	80.05	87

[a]The distributions were obtained by applying Boltzmann statistics at 298.15 K to the entire set of conformers in the gas phase and solution, respectively, in complete analogy to ref 32. [b]Owens et al.[33]

carbohydrate ring exists in a $^4C_1$ chair conformation, while the prolyl side-chain seems to adopt a Cγ-exo conformation (see the Supporting Information for an explanation). There is very little variation in both the bond angles and bond lengths between the two isomers. Thus, the *cis*–*trans* isomerization appears to have little influence on these geometrical parameters. This is a consequence of the stiff molecular framework imposed by the fused rings in *GlcProH*, and it can be contrasted to the much more fexible spirocyclic systems *Ac-Glc3(S)-HypH-OMe* previously investigated.[32]

The backbone dihedral angles and the pseudorotational parameters *A* and *P* of the same two most stable *cis* and *trans* isomers in water are provided in Table 3. The parameters *A* and *P* are determined from the dihedral angles around the five-membered ring χ^j (j = 0–4, Scheme 3), according to the procedure of Hudaky and Perczel.[91] These two parameters characterize the ring-puckering of the five-membered ring, Table 3.

The variation in the backbone dihedral angles is somewhat larger than that for the bond angles. Apart from ω' that obviously has the largest change since it characterizes the isomers as either *cis* or *trans*, there is also a change of some 13° in ϕ and of about 7° in ψ, which likely reflects some accommodation of the prolyl C-terminal carbonyl group upon isomerization. However, the ring puckering parameters, specifically the puckering amplitude (A) and the state of the pucker in the pseudorotational pathway (P), are very similar between the two conformers, Table 3.

Intramolecular Hydrogen Bonding. For the closely related spirocyclic peptide mimic compounds *Ac-Glc3(S)-HypH-OMe* studied previously,[32] Scheme 1, we found that intramolecular hydrogen bonding can play an important role in stabilizing certain conformers. At the same time, however, relatively minor structural differences may have a profound influence on the existence of such hydrogen bonds. For instance, in the case of *Ac-Glc3(S)-HypH-OMe*, we studied two compounds that only differ in the stereochemistry of one carbon center, the C_δ carbon, and intramolecular hydrogen bonds were mostly found in only one stereoisomer. Therefore, internal hydrogen bonding for the most stable conformers of *GlcProH* in water was investigated. As discussed in the Computational Details, a bond length cutoff of 2.5 Å between the hydrogen atom and an adjacent heavy atom was used for this analysis.

Among the 14 most stable isomers in solution (those that contribute at least 0.5% to the total distribution using Boltzmann statistics), only one type of intramolecular hydrogen bond was observed. This hydrogen bond exists between the prolyl N-terminal amide carbonyl oxygen and the sugar 3C hydroxyl group (C_3—OH···O=C′—N), as illustrated in Figure 3. The bond distance between the hydrogen atom and its neighbor is approximately 1.79 Å, indicating a rather strong hydrogen bond. The conformers with this type of hydrogen bond are conformers 5, 8, 10, and 13 (Table S2 of the Supporting Information). The same type of intramolecular hydrogen bond also exists in some of the remaining minor conformers, for instance, conformers 15 and 23. The listed conformers contribute to approximately 19.9% of the total distribution. All of these conformers are in the *cis* conformation with respect to the prolyl amide bond; no intramolecular hydrogen bonds were found for the *trans* isomers. Similar hydrogen bonds exist in the gas phase *cis* conformers also. The intramolecular hydrogen bonds are also illustrated in Figures 1

Table 2. Selected Bond Angles for the Most Stable *cis* and *trans* Isomers of *N-acetyl-GlcProH- NHMe* (*GlcProH*) in Water (deg)[a]

conformer	$\angle N-C^{\alpha}-C^{\beta}$	$\angle C^{\alpha}-C^{\beta}-C^{\gamma}$	$\angle C^{\beta}-C^{\gamma}-C^{\delta}$	$\angle C^{\gamma}-C^{\delta}-N$	$\angle C^{\delta}-N-C^{\alpha}$
t-ε_L[d]	104.0	103.5	103.4	101.4	112.6
c-ε_L[d]	103.2	102.8	104.2	101.7	113.2

[a]See Scheme 3 for the definition of the atoms.

Table 3. Backbone Dihedral Angles (deg), Pseudorotational Parameters A and P[a] (deg), and Dipole Moment (D) for the Most Stable *cis* and *trans* Isomers of *N-acetyl-GlcProH-NHMe* (*GlcProH*) in Water[b]

conformer	ω'	ϕ	ψ	ω	A	P	dipole moment
t-ε_L[d]	171.3	−55.9	143.6	175.6	34.4	−67.7	5.0
c-ε_L[d]	1.9	−69.0	150.7	175.3	33.4	−68.3	10.4

[a]Pseudorotational parameters for the five-membered ring, according to Hudaky and Perczel.[91] A and P are based on the following expression:[91] $\chi^j \cong A\cos(P + 144° j)$, $j = 0, ..., 4$. [b]See Scheme 3 for the definition of the angles.

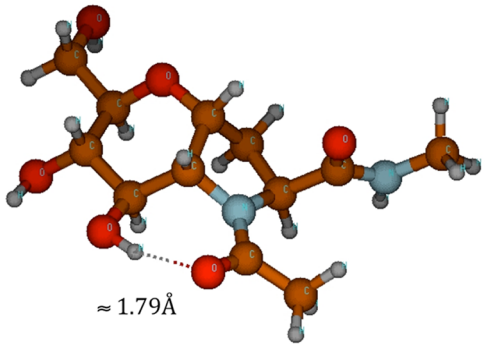


Figure 3. Intramolecular hydrogen bond in the *cis* isomers of *GlcProH*.

and 2, where they are marked as dotted lines (specifically Figures 1a,b,g and 2a). We note that all of the most stable *cis* isomers have this intramolecular hydrogen bond; among the *cis* conformers contributing at least 0.5%, there are none without. Thus, intramolecular hydrogen bonds and the resulting stabilization are at least partly responsible for the high percentage of *cis* conformers in the gas phase.

Visual inspection of the rigid molecular framework reveals the reason for the difference in propensity between *cis* and *trans* isomers with respect to intramolecular hydrogen bonds, Schemes 2 and 3 and Figure 3. The two fused six- and five-membered rings form a relatively rigid structure that only permits a hydrogen bond to the C_3 hydroxyl group in the *cis* conformation. In the *trans* amide conformation, no such hydrogen bond donor is available, and the structure does not greatly change conformation in order to accommodate one. The only suitable hydrogen bond donor, the C-terminal amide NH group, is oriented away from the N-terminal amide carbonyl group. This is in contrast to the previously studied spirocyclic compound *Ac-Glc3(S)-HypH-OMe*, Scheme 1,[32] in which the conformational flexibility around the five- and six-member rings, as well as between the rings, is much greater. Hence, a variety of hydrogen bonds with different hydrogen-bond donors and acceptors are found, which stabilize both *cis* and *trans* isomers in the more flexible analogue.

Solvation Stabilization; Dipole Moment. Above, we have discussed the very different isomer distributions in the gas phase and in aqueous solution, Table 1. The qualitative and quantitative differences in the population distributions between the gas phase and the highly polar solvent (water) can be attributed to different free energies of solvation ($\Delta\Delta G$ of solvation for the different conformers).

Oftentimes, a qualitative understanding of differential free energies of solvation can be achieved using simple arguments based on molecular charge or (for neutral molecules) dipole moment in combination with molecular size.[92−94] For neutral molecules, the free energy of solvation is, to first order, proportional to the dipole moment of the solute squared and inversely proportional to the distance between dipole and polarizable solvent cubed.[65,88−90] In Table 3, we provide the calculated dipole moments for the most stable *cis* and *trans* isomers of *GlcProH* in water. The *cis* conformer has a larger dipole moment compared to the *trans* isomer. Thus, the above-mentioned simple first-order arguments would result in the conclusion of preferential solvation stabilization for the *cis* isomer, in contrast to the actual situation (Table 1). We conclude that these types of arguments are not applicable in the present case. Clearly, higher-order effects are the determining factors, which are accurately captured by continuum solvation models such as PCM, as seen in Table 1.

Upon closer consideration, it is not really surprising that a simple qualitative model that is solely based on the dipole moment fails to properly describe solvation in a polar solvent. Carbohydrate-templated peptide mimics such as the *GlcProH* molecule possess numerous polar hydroxyl groups. In *GlcProH*, these comprise two carboxylic groups, three terminal alcohols, two amide groups, and so on, Schemes 2 and 3. These groups are distributed over various regions in the molecule. Hence, the interactions of the various local bond dipoles with adjacent parts of the polar solvent cannot be properly described by a simple global molecular dipole interacting with the solvent, and proper continuum solvation models or explicit solvent treatment are required. (In some cases, one might have to go even further and include explicit solvation as well.[95])

■ CONCLUSION

In this paper, we have used modern computational chemistry to study the conformational distribution of *N-acetyl-GlcProH-NHMe* (*GlcProH*) in the gas phase and in water. A reliable computational protocol, consisting of systematic and Monte Carlo molecular mechanics searches, followed by accurate DFT calculations that include solvation effects using a continuum solvation model, has been used. This approach ensures that all relevant conformers are included in the analysis according to accurate relative free energies. This computational protocol, as previously developed for flexible analogues, also applies to this more rigid system that possesses two fused rings. Thus, it should be generally applicable to other flexible organic molecules with and without rings.

dx.doi.org/10.1021/jp310690c | J. Phys. Chem. B 2013, 117, 199−205

Solvation is found to be crucial, since the calculated prolyl *cis*–*trans* distributions are essentially opposite in the gas phase (95% *cis*) and solution (80% *trans*). This drastic change is due to differences in the stabilization of different conformers within a polar solvent (free energy of solvation). This effect cannot be explained by simplistic arguments that rely only on the overall dipole moment of the molecule due to the presence of numerous polar bonds (local bond dipoles). However, continuum solvation models, such as PCM, are capable of determining accurate, relative free energies of solvation for different conformers of these analogues. Additionally, there are significant structural changes between gas-phase and solution-optimized structures, emphasizing the need for optimizations in the presence of the polar solvent, as opposed to PCM single-point calculations.

Considering the most stable solution-optimized conformers, there is relatively little change in the overall geometries in going from *cis* to *trans*. This is due to the stiff molecular framework of the fused rings in the *GlcProH* molecule, and it can be contrasted to more flexible systems such as the closely related spirocyclic *Ac-Glc3(S)-HypH-OMe* system, Scheme 1.[32] However, only the *cis* isomers of *GlcProH* possess intramolecular hydrogen bonds, and the resulting stabilization at least partially explains the large excess of *cis* in the gas phase. Because of the relatively stiff molecular framework imposed by the fused five- and six-membered rings, only *cis* isomers possess a geometry that brings hydrogen bond donors and acceptors into close enough proximity. This can be contrasted to the spirocyclic *Ac-Glc3(S)-HypH-OMe* system, Scheme 1, that is much more flexible and therefore shows a much greater variety in the types and strengths of intramolecular hydrogen bonds. In this context, it is noteworthy that hydrogen bonds in related hydroxyproline analogues have been show to play an important role in the control of prolyl *cis/trans* isomerization in larger peptide systems.[96] Further studies could include modeling the effect of hydroxyl substitution on the *cis*–*trans* equilibrium in the absence of intramolecular hydrogen bonding. Moreover, when compared to normal proline, the *cis/trans* rotamer population is greatly affected by the presence of the C-glucosylated proline ring. For instance, the NMR-determined *cis* ratio of AcN-acetyl-Pro-NHMe is 28%,[97] which is significantly higher when compared to N-acetyl-GlcProH-NHMe (13%), indicating that the presence of the fused glucose ring stabilizes the *trans* rotamer population in this proline analogue. The *trans* prolyl amide conformer plays an important function in structural proteins such as collagen and the polyproline helix.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Full citations for refs 36, 48, and 78; description of the nomenclature as per ref 32; Tables S1 and S2 containing geometry and energetic details of the most stable conformers in the gas phase and aqueous solution, respectively; Table S3 showing selected bond lengths. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: schrecke@cc.umanitoba.ca (G.S.); schweize@cc.umanitoba.ca (F.S.).

Present Address

[‖]Institut Européen de Chimie et Biologie, Université de Bordeaux—CNRS UMR 5248, Pessac, 33607, France.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canada Foundation for Innovation (CFI). S.D.W. also thanks the Canada Research Chairs program.

■ REFERENCES

(1) Berg, J.; Tymoczko, J. L.; Stryer, L. *Biochemistry*, 6th ed.; W. H. Freeman & Co.: New York, 2006.
(2) Wilmot, C. M.; Thornton, J. M. *J. Mol. Biol.* **1988**, *203*, 221–232.
(3) Fischer, G.; Schmid, F. X. *Biochemistry* **1990**, *29*, 2205–2212.
(4) Mortensen, M.; Husmann, R.; Veri, E.; Bolm, C. *Chem. Soc. Rev.* **2009**, *38*, 1002–1010.
(5) Che, Y.; Marshall, G. R. *J. Org. Chem.* **2004**, *69*, 9030–9042.
(6) Kang, Y. K.; Byun, B. J. *J. Phys. Chem. B* **2007**, *111*, 5377–5385.
(7) Delaney, N. G.; Madison, V. *J. Am. Chem. Soc.* **1982**, *104*, 6635–6641.
(8) Flippen-Anderson, J. L.; Gilardi, R.; Karle, I. L.; Frey, M. H.; Opella, S. J.; Gierasch, L. M.; Goodman, M.; Madison, V.; Delaney, N. G. *J. Am. Chem. Soc.* **1983**, *105*, 6609–6614.
(9) Samanen, J.; Zuber, G.; Bean, J.; Eggleston, D.; Romoff, T.; Kopple, K.; Saunders, M.; Regoli, D. *Int. J. Pept. Protein Res.* **1990**, *35*, 501–509.
(10) Beausoleil, E.; Lubell, W. D. *J. Am. Chem. Soc.* **1996**, *118*, 12902–12908.
(11) Sharma, R.; Lubell, W. D. *J. Org. Chem.* **1996**, *61*, 202–209.
(12) Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T. *Nature* **1998**, *392*, 666–667.
(13) DeRider, M. L.; Wilkens, S. J.; Waddell, M. J.; Bretscher, L. E.; Weinhold, F.; Raines, R. T.; Markley, J. L. *J. Am. Chem. Soc.* **2002**, *124*, 2497–2505.
(14) Quancard, J.; Labonne, A.; Jacquot, Y.; Chassaing, G.; Lavielle, S.; Karoyan, P. *J. Org. Chem.* **2004**, *69*, 7940–7948.
(15) Shoulders, M. D.; Hodges, J. A.; Raines, R. T. *J. Am. Chem. Soc.* **2006**, *128*, 8112–8113.
(16) Cadamuro, S. A.; Reichold, R.; Kusebauch, U.; Musiol, H. J.; Renner, C.; Tavan, P.; Moroder, L. *Angew. Chem., Int. Ed.* **2008**, *47*, 2143–2146.
(17) Kang, Y. K.; Byun, B. J.; Park, H. S. *Biopolymers* **2011**, *95*, 51–61.
(18) Kern, D.; Schutkowski, M.; Drakenberg, T. *J. Am. Chem. Soc.* **1997**, *119*, 8403–8408.
(19) Keller, M.; Sager, C.; Dumy, P.; Schutkowski, M.; Fischer, G. S.; Mutter, M. *J. Am. Chem. Soc.* **1998**, *120*, 2714–2720.
(20) Tam, J. P.; Miao, Z. W. *J. Am. Chem. Soc.* **1999**, *121*, 9013–9022.
(21) Kang, Y. K.; Park, H. S. *J. Phys. Chem. B* **2007**, *111*, 12551–12562.
(22) Chaume, G.; Barbeau, O.; Lesot, P.; Brigaud, T. *J. Org. Chem.* **2010**, *75*, 4135–4145.
(23) Galardy, R. E.; Alger, J. R.; Liakopoulou-Kyriakides, M. *Int. J. Pept. Protein Res.* **1982**, *19*, 123–132.
(24) Jhon, J. S.; Kang, Y. K. *J. Phys. Chem. B* **2007**, *111*, 3496–3507.
(25) Wagaw, S.; Rennels, R. A.; Buchwald, S. L. *J. Am. Chem. Soc.* **1997**, *119*, 8451–8458.
(26) Viswanathan, R.; Prabhakaran, E. N.; Plotkin, M. A.; Johnston, J. N. *J. Am. Chem. Soc.* **2003**, *125*, 163–168.
(27) Koep, S.; Gais, H. J.; Raabe, G. *J. Am. Chem. Soc.* **2003**, *125*, 13243–13251.
(28) Jeannotte, G.; Lubell, W. D. *J. Am. Chem. Soc.* **2004**, *126*, 14334–14335.

(29) Jeannotte, G.; Lubell, W. D. *J. Org. Chem.* **2004**, *69*, 4656−4662.
(30) Zhang, K.; Teklebrhan, R. B.; Schreckenbach, G.; Wetmore, S.; Schweizer, F. *J. Org. Chem.* **2009**, *74*, 3735−3743.
(31) Zhang, K. D.; Schweizer, F. *Synlett* **2005**, No. 20, 3111−3115.
(32) Teklebrhan, R. B.; Zhang, K.; Schreckenbach, G.; Schweizer, F.; Wetmore, S. D. *J. Phys. Chem. B* **2010**, *114*, 11594−11602.
(33) Owens, N. W.; Braun, C.; Schweizer, F. *J. Org. Chem.* **2007**, *72*, 4635−4643.
(34) Manzoni, L.; Colombo, M.; May, E.; Scolastico, C. *Tetrahedron* **2001**, *57*, 249−255.
(35) Petter, R. C. *Tetrahedron Lett.* **1989**, *30*, 399−402.
(36) Blankley, C. J.; Kaltenbronn, J. S.; Dejohn, D. E.; Werner, A.; Bennett, L. R.; Bobowski, G.; Krolls, U.; Johnson, D. R.; Pearlman, W. M.; Hoefle, M. L.; et al. *J. Med. Chem.* **1987**, *30*, 992−998.
(37) Li, W. H.; Moeller, K. D. *J. Am. Chem. Soc.* **1996**, *118*, 10106−10112.
(38) Chakraborty, T. K.; Jayaprakash, S.; Diwan, P. V.; Nagaraj, R.; Jampani, S. R. B.; Kunwar, A. C. *J. Am. Chem. Soc.* **1998**, *120*, 12962−12963.
(39) Grotenbreg, G. M.; Timmer, M. S. M.; Llamas-Saiz, A. L.; Verdoes, M.; van der Marel, G. A.; van Raaij, M. J.; Overkleeft, H. S.; Overhand, M. *J. Am. Chem. Soc.* **2004**, *126*, 3444−3446.
(40) Chakraborty, T. K.; Srinivasu, P.; Rao, R. V.; Kumar, S. K.; Kunwar, A. C. *J. Org. Chem.* **2004**, *69*, 7399−7402.
(41) Maigret, B.; Perahia, D.; Pullman, B. *J. Theor. Biol.* **1970**, *29*, 275−291.
(42) Pullman, B.; Maigret, B.; Perahia, D. *Theor. Chim. Acta* **1970**, *18*, 44−56.
(43) Farmer, B. L.; Hopfinger, A. J. *Macromolecules* **1974**, *7*, 793−796.
(44) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(45) Fischer, S.; Dunbrack, R. L.; Karplus, M. *J. Am. Chem. Soc.* **1994**, *116*, 11931−11937.
(46) Kang, Y. K. *J. Phys. Chem.* **1996**, *100*, 11589−11595.
(47) McDonald, D. Q.; Still, W. C. *J. Org. Chem.* **1996**, *61*, 1385−1391.
(48) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(49) Jhon, J. S.; Kang, Y. K. *J. Phys. Chem. A* **1999**, *103*, 5436−5439.
(50) Improta, R.; Benzi, C.; Barone, V. *J. Am. Chem. Soc.* **2001**, *123*, 12568−12577.
(51) Benzi, C.; Improta, R.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2002**, *23*, 341−350.
(52) Rankin, K. N.; Boyd, R. J. *J. Phys. Chem. A* **2002**, *106*, 11168−11172.
(53) Kang, Y. K.; Choi, H. Y. *Biophys. Chem.* **2004**, *111*, 135−142.
(54) Enriz, R. D.; Morales, M. E.; Baldoni, H. A. *J. Mol. Struct.: THEOCHEM* **2005**, *731*, 177−185.
(55) Sahai, M. A.; Kehoe, T. A. K.; Koo, J. C. P.; Setiadi, D. H.; Chass, G. A.; Viskolcz, B.; Penke, B.; Pai, E. F.; Csizmadia, I. G. *J. Phys. Chem. A* **2005**, *109*, 2660−2679.
(56) Song, I. K.; Kang, Y. K. *J. Phys. Chem. B* **2006**, *110*, 1915−1927.
(57) Kang, Y. K. *J. Phys. Chem. B* **2006**, *110*, 21338−21348.
(58) Che, Y.; Marshall, G. R. *Biopolymers* **2006**, *81*, 392−406.
(59) Aliev, A. E.; Courtier-Murias, D. *J. Phys. Chem. B* **2007**, *111*, 14034.
(60) Flores-Ortega, A.; Jimenez, A. I.; Cativiela, C.; Nussinov, R.; Aleman, C.; Casanovas, J. *J. Org. Chem.* **2008**, *73*, 3418.
(61) Flores-Ortega, A.; Casanovas, J.; Nussinov, R.; Aleman, C. *J. Phys. Chem. B* **2008**, *112*, 14045.
(62) Aliev, A. E.; Bhandal, S.; Courtier-Murias, D. *J. Phys. Chem. A* **2009**, *113*, 10858−10865.
(63) Kang, Y. K.; Park, H. S. *Biopolymers* **2009**, *92*, 387−398.
(64) Kang, Y. K.; Byun, B. J. *J. Comput. Chem.* **2010**, *31*, 2915−2923.
(65) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*, 2nd ed.; Wiley: New York, 2004.
(66) Schreckenbach, G.; Shamov, G. A. *Acc. Chem. Res.* **2010**, *43*, 19−29.
(67) Spartan '02; Build 119; Wavefunction, Inc.: Irvine, CA, 2002.
(68) Hehre, W. J. *A Guide to Molecular Mechanics and Quantum Chemical Calculations*; Wavefunction, Inc.: Irvine, CA, 2003.
(69) Leach, A. R.; Gillet, V. J. *An Introduction to Chemoinformatics*; Springer: New York, 2003.
(70) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490−519.
(71) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 520−552.
(72) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 553−586.
(73) Halgren, T. A.; Nachbar, R. B. *J. Comput. Chem.* **1996**, *17*, 587−615.
(74) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 616−641.
(75) Halgren, T. A. *J. Comput. Chem.* **1999**, *20*, 720−729.
(76) Halgren, T. A. *J. Comput. Chem.* **1999**, *20*, 730−748.
(77) Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*; Wiley Verlag Chemie: New York, 2000.
(78) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; et al. *Gaussian 03*; Gaussian, Inc.: Wallingford, CT, 2004.
(79) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(80) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B: Condens. Matter* **1988**, *37*, 785−789.
(81) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623−11627.
(82) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117−129.
(83) Flores-Ortega, A.; Casanovas, J.; Assfeld, X.; Aleman, C. *J. Org. Chem.* **2009**, *74*, 3101−3108.
(84) Morozov, A. V.; Kortemme, T.; Tsemekhman, K.; Baker, D. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 6946−6951.
(85) Perczel, A.; McAllister, M. A.; Csaszar, P.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1993**, *115*, 4849−4858.
(86) Perczel, A.; McAllister, M. A.; Csaszar, P.; Csizmadia, I. G. *Can. J. Chem.* **1994**, *72*, 2050−2070.
(87) Song, I. K.; Kang, Y. K. *J. Phys. Chem. B* **2005**, *109*, 16982−16987.
(88) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027−2094.
(89) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161−2200.
(90) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999−3093.
(91) Hudaky, I.; Perczel, A. *J. Mol. Struct.: THEOCHEM* **2003**, *630*, 135−140.
(92) Shamov, G. A.; Schreckenbach, G.; Martin, R. L.; Hay, P. J. *Inorg. Chem.* **2008**, *47*, 1465−1475.
(93) Pan, Q.-J.; Schreckenbach, G. *Inorg. Chem.* **2010**, *49*, 6509−6517.
(94) Asaduzzaman, A. M.; Schreckenbach, G. *Phys. Chem. Chem. Phys.* **2011**, *13*, 15148−15157.
(95) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **2006**, *110*, 2493−2499.
(96) Kuemin, M.; Nagel, Y. A.; Schweizer, S.; Monnard, F. W.; Ochsenfeld, C.; Wennemers, H. *Angew. Chem., Int. Ed.* **2010**, *49*, 6324−6327.
(97) Beausoleil, E.; Sharma, R.; Michnick, S. W.; Lubell, W. D. *J. Org. Chem.* **1998**, *63*, 6572−6578.