CHEMIJA. 2007. Vol. 18. No. 3. P. 1–6

The study of the vibrational Raman spectra modeling of solvated cysteamine

Laurynas Riauba

Department of Organic Chemistry, Vilnius University, Naugarduko 24, LT-03225 Vilnius, Lithuania
E-mail: laurynas.riauba@chf.vu.lt

The characteristics of Raman spectra calculations of solvated molecular structures were investigated. Cysteamine hydrochloride solution in water was chosen as the model system. The calculations were performed using an implicit solvation model Integral Equation Formalism Polarizable Continuum Model (IEFPCM) with explicit solvent molecules added to the structure model. Geometry optimization of such structures might be time consuming due to a flat energy surface. The analysis of alternation of the calculated Raman spectra during the geometry optimization process revealed that a considerably lower number of optimization steps was necessary to get semi-accurate results. We have found that 10–15 optimization steps are enough for such models, and calculation errors for such models do not exceed 5–10%. The correlations between the calculation errors and internal forces on atoms or changes of the geometric parameters of the models are observed.

Key words: Raman spectra, geometry optimization, solvation, DFT, cysteamine

INTRODUCTION

Investigating the fine structures of solvated molecules is a topic of considerable importance since most chemical and biochemical processes occur in solutions. Computational modelling of such species is a common tool for structural studies. The models of solvated molecules may provide some fundamental information about molecular systems; therefore, there is great demand for accurate and straightforward theoretical models that account for the effects of solvation. The most common and computationally efficient way to treat a solvent is to use continuum solvation models.

The history of and theory of continuum solvation models have been laid out extensively in various articles [1–4]. The most prevailing of them are polarized continuum models (PCM) [1.5–7]. Three main different approaches exist to exercise PCM calculations, namely dielectric PCM (D-PCM), the second model which treats a surrounding medium as a conductor instead of dielectric (C-PCM) and, finally, integral equation formalism implementation of PCM (IEF-PCM) [8–9]. Very similar to PCM is a conductor-like screening model (COSMO) [10–13], which is also quite current and widely applicable in software packages.

Usually implicit solvation models are created and optimized to perform the calculations of the solvated molecule energy. But there are certain cases when the calculation results can be improved by the use of some explicit solvent molecules [14-15]. These explicit solvent molecules may provide some fundamental information [16–17] on the interactions between the solvent and solute molecules. For typical *ab-initio* calculations, an inclusion of many solvent molecules is not appropriate due to huge computational requirements. Therefore, it is common to use simpler models with solvated molecule including few solvent molecules (it is necessary to choose an appropriate number of solvent molecules, which could provide information on the interactions). Nonetheless, it should be noted that even such approach has a hidden computational time. Such models contain more or less weakly connected molecules. Usually these molecules are bound by a hydrogen bonding. Hydrogen bonds are among the strongest noncovalent bonds, but nevertheless by orders of magnitude they are weaker than covalent bonds. Therefore, the energy surface of such molecular structures near the optimized structures is very flat. Most optimization algorithms work very well when the energy gradient is moderate, but may behave worse when the energy surface is flat. A typical optimization pattern of such weakly connected systems is fast structure convergence of separate molecules and slow relative alignment of them. In the worst case, some structures do not get optimized at all due to slight vibrations of the solvent molecules forwards and backwards.

Some computational tasks require good geometry convergence to perform the final calculations. An example of such task is a calculation of vibrational infrared absorption or Raman spectra. It is a usual recommendation for the vibrational analysis to find fully converged geometry. It is stated that spectra and energy depend on a precise optimized structure. This recommendation is so strict that even structure optimization and vibrational analysis should be performed using the same basis set. Indeed, the physical basis for the vibrational analysis assumes that energy derivatives in respect of the coordinates are equal to zero, and this condition is fulfilled at energy surface minimum. If geometry is not optimized, the vibrational analysis usually yields one or more imaginary frequencies. These strict requirements for geometry are full of sense when modelling molecules in vacuo.

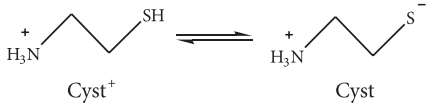
Modelling of the solvated molecules have some more aspects. The usage of explicit or implicit solvation models bring errors by themselves. The implicit models contain some correlated values in the calculation scheme. The explicit solvation models usually take into account only few solvent molecules; therefore, many solvent–solvent molecular interactions are neglected. It should be pointed out that both types of solvation models neglect the dynamic nature of the solvation process.

Thus, the calculations of the vibrational spectra, which need accurate geometries, depend on the solvation model. In this work we are analyzing the calculated Raman spectra during the optimization process. We have chosen cysteamine hydrochloride solutions in water as a model system. Cysteamine (2-aminoethanethiol) is a small bifunctional molecule widely used as a chelating ligand in coordination chemistry, biochemistry, and for construction of functional self-assembled monolayers (SAMs) on metal surfaces [18]. Its SH group forms a strong metal–sulfur chemical bond upon the adsorption on gold, silver and copper surfaces, while the NH_2 group interacts with ions [19], organic molecules [20] and proteins [21–22] in the solution phase used to prepare molecular structures with specific properties. Ionization and solvation of the amine group in aqueous solutions controls the binding properties of the cysteamine linkage to the solution phase species. The Raman spectra of cysteamine hydrochloride in solution were previously reported and the bands were assigned [23–26].

In this work we are demonstrating that during the calculations of Raman spectra of solvated cysteamine molecules, a partial optimization with almost converged geometries is sufficient. For the molecular systems studied herein, this could reduce the calculation time-span by five times and even more for larger systems.

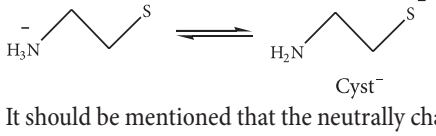
METHODS

All calculations were performed with the software package Gaussian for Windows G03W rev D03 [27]. All the calculations were performed using DFT method with the use of B3LYP hybrid functional. The 6-31++G(d,p) gaussian basis sets were used. Cysteamine has three ionic forms with formal charges +1, 0 and –1, denoted as $Cyst^+$, Cyst and $Cyst^-$ respectively, which are in equilibrium according to equations 1 and 2.



Cyst$^+$ Cyst (1)

and



Cyst$^-$ (2)

It should be mentioned that the neutrally charged form Cyst with zero formal charge actually is zwitterionic in water solutions. The models for all the three ionic forms were generated. These models additionally included three explicit water molecules, positioned randomly at initial geometry.

All the vibrational spectra from intermediate optimization steps were calculated using the same calculation parameters. The convergence of geometry during the optimization was calculated according to the following equation:

$$\Delta\bar{x}^m = \sqrt{\frac{\sum_{n=1..N}(x_n^m - x_n^f)^2 + (y_n^m - y_n^f)^2 + (z_n^m - z_n^f)^2}{N}}, \quad (3)$$

where **m** is geometry at m-*th* step, **N** is the number of atoms, $\mathbf{x}_\mathbf{n}^\mathbf{f}$ denotes the coordinate "x" of atom n at the final geometry.

The mean relative errors of the calculated spectra were calculated using the following equations:

$$\Delta\bar{\nu}^m = \frac{\sum_{n=1..K}\left|\nu_n^m - \nu_n^f\right|}{K} \quad (4)$$

and

$$\Delta\bar{I}^m = \frac{\sum_{n=1..K}\left|\left(\frac{I_n^m}{I_n^f} - 1\right)\cdot 100\%\right|}{K} \quad (5)$$

RESULTS AND DISCUSSION

For this study we have selected a cysteamine molecule, the ionization of which was examined earlier by the spectroscopic and computational methods [23]. Herein we demonstrate that the optimization of the cysteamine ionic forms generally takes different numbers of optimization cycles. It was found that the main changes of geometries evolved during the initial 10–15 cycles, further geometries were changing more slowly. Fig. 1 shows a progress of geometry convergence and relative changes of geometry during the optimization process. Some models converge to the final geometry monotonically, like the neutral Cyst and the positive $Cyst^+$ species model, while others may fluctuate, like the negative form $Cyst^-$ model. A more detailed analysis shows that these fluctuations are mainly caused by the "vibrations" of water molecules. It can be noticed as well that the geometries are changing at different rates, and there are some intervals of fast changes as well as intervals of relatively stable structures. The intervals of rapid changes can be naturally associated with forces on the atoms. This fact is not unexpected considering that geometry optimization algorithms use the force information when updating geometries.

The goal of the geometry optimization is to obtain the structure with minimal energy. Usually the objective is to find a global minimum, but sometimes algorithms have to "climb" up to higher energies on the energy surface to leave a local minimum and find the global minimum. These optimization profiles can be seen in the optimization energy graphs. Fig. 2 demonstrates an energy calculation error change of cysteamine forms during the optimization. The energies of the initial structures differ from the optimized structure energies by 10–14 kJ/mol. During the optimization process, the electronic energies of all ionic forms monotonically decrease to the final values, and the energy calculation error is always positive and converges to zero. As mentioned above, theoretically it is possible that the energy may rise to a certain local maximum.

However, the calculated energies with zero-point corrections or calculated free energies behave more erratically. Nevertheless, usually the calculation error does not exceed 5–10 kJ/mol after the

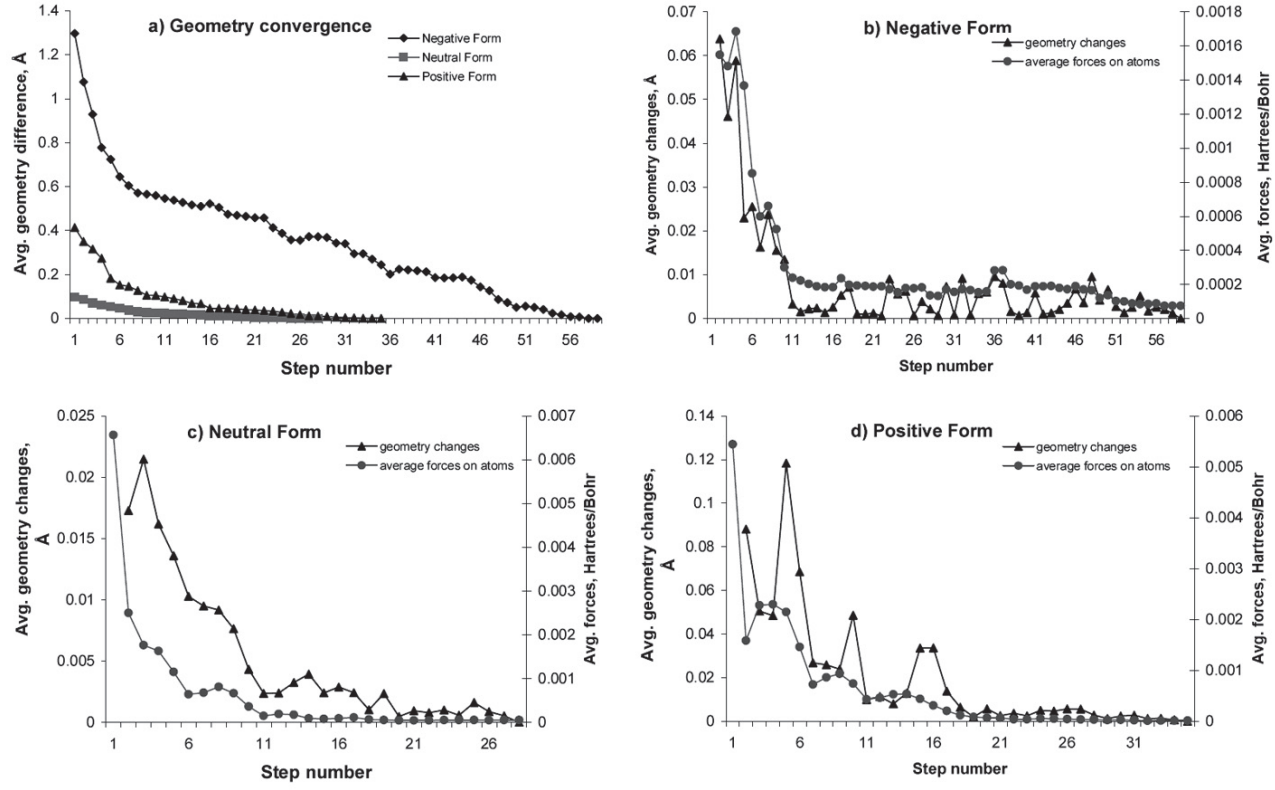


Fig. 1. Overall average geometry convergence (a) and relative average changes in the geometry (in Å) and forces (in Hartrees/Bohr) in negative $Cyst^{-}$(b), neutral Cyst (c) and positive $Cyst^{+}$ (d) forms of cysteamine

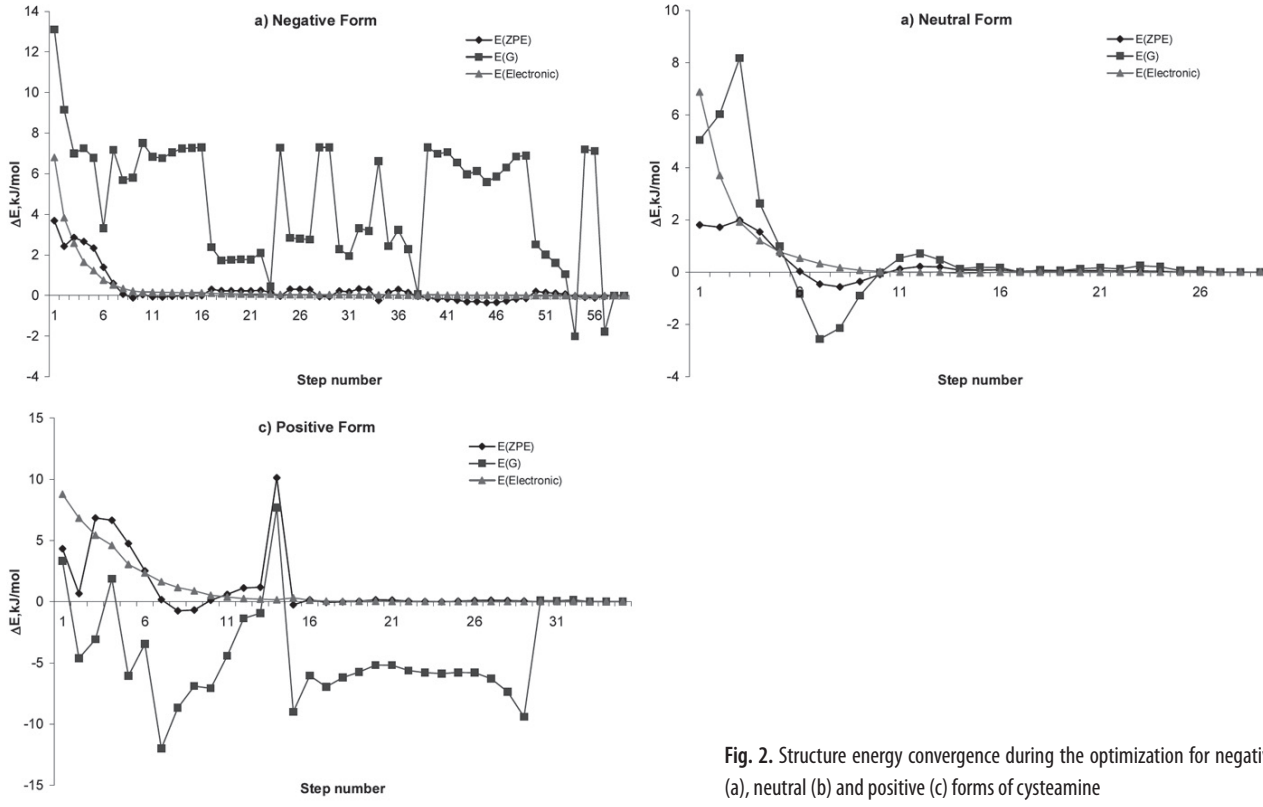


Fig. 2. Structure energy convergence during the optimization for negative (a), neutral (b) and positive (c) forms of cysteamine

first few optimization cycles. Some authors suggest that the usage of polarizable continuum models contain implicit calculation of correction factors, therefore, the calculation of zero-point energy and subsequently free energy is not correct. An interesting relationship between free energy and minimal frequency (or maximal imaginary frequency) is observed for the negative form of the cysteamine molecule. There are two clearly visible correlations for geometries with imaginary frequencies and with all positive frequencies.

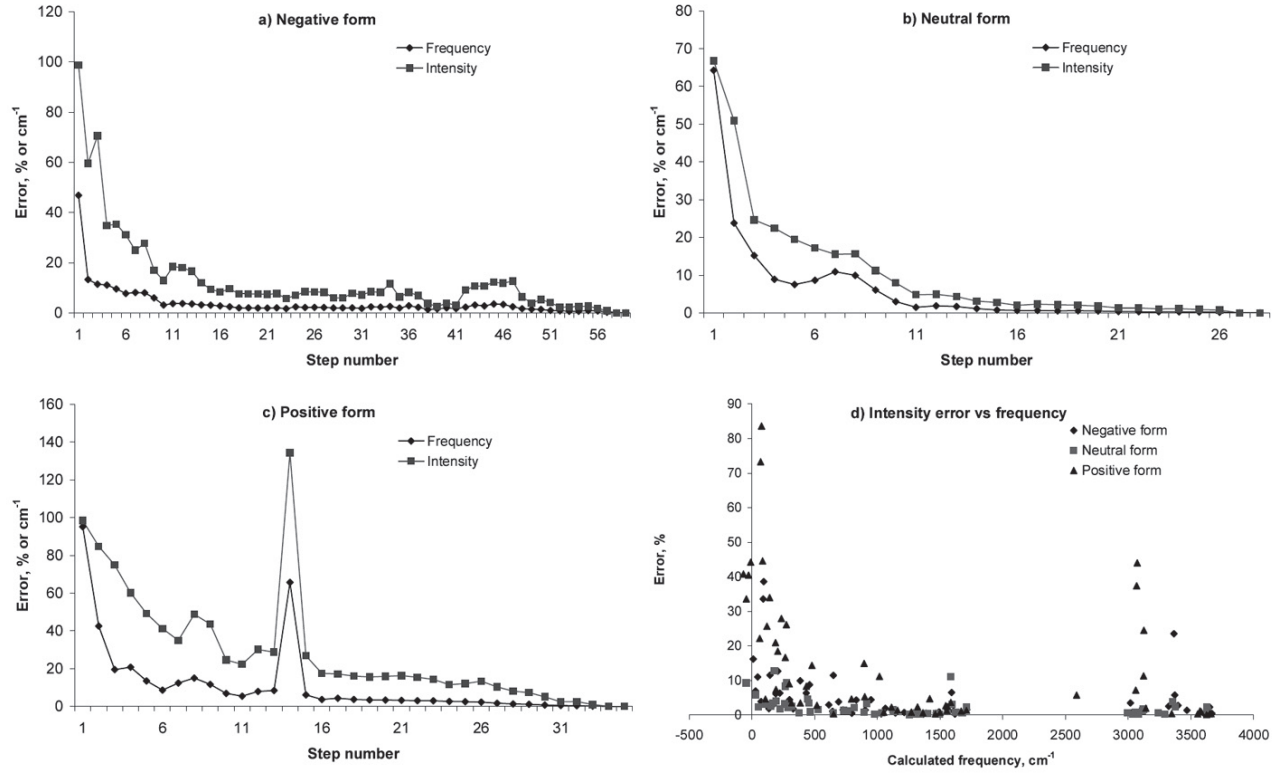


Fig. 3. Frequency and intensity errors during the optimization for negative (a), neutral (b), positive (c) forms of cysteamine and the intensity error calculation dependence on the calculated frequency (d)

The main goal of this work was to analyze the changes of the calculated Raman spectra during the optimization. The experimentally recorded Raman spectrum is continuous, but it can be simulated using calculated spectral band frequencies and intensities. Therefore, it is possible to separate the analysis of peak frequencies and peak intensity. Fig. 3 (a)–(c) shows the frequency and intensity error changes during the optimization. Generally, the errors have the highest values during the first optimization cycles and later they asymptotically approximate to zero. While the general tendency is to decrease asymptotically to zero, the calculation error graphs have some regions of steep descending, usually during the beginning of the optimization, and some local "hills" on the graphs. Probably these can be associated with less stable transition geometries. It should be noted that on average the errors are quite low. There are some exceptions to this general tendency. For example, the positive form $Cyst^+$ has quite large errors at 13–14 optimization cycles. But it turns out that this phenomenon is caused by saddle point geometry – the highest calculated imaginary frequency reaches –2000 cm^{-1} and this indicates very unstable transition geometry. In general, the errors reach acceptable levels after the first 5–15 cycles, but the geometry optimization may take much more cycles to find the converged geometry. After these initial optimization steps, the average absolute frequency calculation error remains below 10–20 cm^{-1} and the average relative intensity error is below 20%. Compared to experimental data [23], similar or slightly higher errors can be seen when comparing the experimental and calculated spectra.

The main changes to the model geometry are done on water molecules. Therefore, it is logical to expect main optimization errors on vibrations, which are associated with simultaneous movements of water and cysteamine atom. The diagram in Fig. 3 (d) shows the intensity error dependence on the calculated frequency. Three zones of higher errors can be distinguished. The first one is below 500 cm^{-1}, the second one is at ca 1600cm^{-1} and the third one is above 3200 cm^{-1}. All these zones can be associated with vibrational water frequencies – rotational, deformational and stretching. Except these zones of higher possibility of the calculation errors, the average calculation error during the optimization is usually below 10%.

It can be concluded from the above figures that the energy, geometry, frequency error as well as the intensity error curves have similar profiles, all of them having exponential decay function shape. Table contains data on the correlation between the frequency and intensity errors versus cysteamine structure energy, geometry and internal forces.

It can also be concluded that there is a clear correlation between the average forces on atoms and geometry changes. The correlation coefficients are vary from 0.86 for the positive form of cysteamine Cys^+ to 0.95 for the neutral form Cys. A strong relationship can be observed for the energy calculation error and average forces on atoms – correlation coefficients vary from 0.9 to 0.98. It should be noted that a good correlation between frequency and intensity calculation errors can also be observed. This observation does not surprise, because the calculation of frequencies and optical Raman intensities both depend on the second derivatives of energy with respect to coordinates.

Table. **Correlations between Raman spectra band frequency (Δν) and intensity (ΔI) calculation errors, calculated energies (E), mean geometry changes (Δx) and mean forces on atoms (F)**

Cysteamine form		Δν	ΔI	E	Δx(change)	Δx(final)	F
Negative	Δν	1.00	0.91	0.95	0.89	0.74	0.95
	ΔI	0.91	1.00	0.94	0.86	0.84	0.91
	E	0.95	0.94	1.00	0.91	0.75	0.90
	Δx(change)	0.89	0.86	0.91	1.00	0.69	0.93
	Δx(final)	0.74	0.84	0.75	0.69	1.00	0.75
	F	0.95	0.91	0.90	0.93	0.75	1.00
Neutral	Δν	1.00	0.94	0.97	0.92	0.83	0.99
	ΔI	0.94	1.00	0.95	0.89	0.95	0.94
	E	0.97	0.95	1.00	0.81	0.84	0.98
	Δx(change)	0.92	0.89	0.81	1.00	0.96	0.95
	Δx(final)	0.83	0.95	0.84	0.96	1.00	0.85
	F	0.99	0.94	0.98	0.95	0.85	1.00
Positive	Δν	1.00	0.88	0.74	0.44	0.73	0.79
	ΔI	0.88	1.00	0.70	0.56	0.76	0.70
	E	0.74	0.70	1.00	0.78	0.97	0.92
	Δx(change)	0.44	0.56	0.78	1.00	0.79	0.86
	Δx(final)	0.73	0.76	0.97	0.79	1.00	0.91
	F	0.79	0.70	0.92	0.86	0.91	1.00

CONCLUSIONS

Structure optimization is a required step for the calculations of vibrational frequencies and intensities of IR and Raman spectra peaks. It is usually stated that correct computations should involve structural optimizations to full convergence. However, loose molecular systems having solvent molecules are difficult to optimize. We have found that 10–15 optimization steps are enough for semi-accurate results and the calculation errors for our models do not exceed 5–10%. Usually the highest calculated spectral band intensity errors are either at low (<500cm^{-1}) or at high frequencies (>3200cm^{-1}).

It should be noted that there is some correlation between the calculation errors and the internal forces on atoms or changes of the model geometries.

Received 28 August 2007 25
Accepted 05 September 2007

References

1. J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, **105**, 2999 (2005).
2. J. L. Rivail and D. Rinaldi. In: J. Leszczynski (ed.), *Computational Chemistry Review of Curent Trends*, p. 139, World Scientific, New York (1995).
3. C. J. Cramer and D. G. Truhlar. In: K. B. Lipkowitz and D. B. Boyd (eds.), *Reviews in Computational Chemistry*, Vol 6, VCH, New York (1995).
4. C. J. Cramer and D. G. Truhlar. In: O. Tapia and J. Betran (eds), *Solvent Effects and Chemical Reactivity*, Kluwer, Dodrecht (1996).
5. G. A. Bagyan, A. K. Valeev, I. K. Koroleva and N. V. Soroka Russ, *J. Inorg. Chem.*, **28**, 1142 (1983).
6. R. Cammi and J. Tomasi, *J. Comput. Chem.*, **16**, 1449 (1995).
7. C. Amovilli, V. Barone, R. Cammi, E. Cance`s, M. Cossi, B. Menucci and C. S. Pommelli, *Adv. Quantum. Chem.,* **32**, 227 (1998).
8. E. Cance`s, B. Mennucci and J. Tomasi, *J. Chem. Phys.,* **107**, 3032 (1997).
9. B. Mennucci, E. Cance`s, J. Tomasi, *J. Phys. Chem. B.*, **101**, 10506 (1997).
10. A. Klamt and G. Schüürmann, *J. Chem. Soc. Perkin Trans.*, **2**, 799 (1993).
11. A. Klamt and V. Jonas, *J. Chem. Phys.*, **105**, 9972 (1996).
12. T. N. Truong and E.V. Stefanovich, *J. Chem. Phys.*, **103**, 3709 (1995).
13. T. N. Truong, *Int. Rev. Phys. Chem.*, **17**, 525 (1998).
14. B. Pullman, S. Miertius and D. Perahia, *Theor. Chim. Acta*, **50**, 317 (1979).
15. B. Mennucci and J. M. Martinez, *J. Phys. Chem. B.*, **109**, 9818 (2005).
16. A. Warshel and R. M Weiss, *J. Am. Chem. Soc.*, **102**, 6218 (1980).
17. J. Kubelka, R. Huang and T. A. Keiderling, *J. Phys. Chem. B*, **109**, 8231 (2005).
18. F. Arias, L. A. Godinez, S. R. Wilson, A. E. Kaifer and L. Echegoyen, *J. Am. Chem. Soc.,* **118**, 6086 (1996).

19. A. Michota, A. Kudelski and J. Bukowska, *J. Raman Sectrosc.*, **32**, 345 (2001).
20. E. Katz and H. L. Schmidt, *J. Electroanal. Chem*., **360**, 337 (1993).
21. L. S. Wong, V. L. Vilker, W. T. Yap and V. Reipa, *Langmuir*, **11**, 4818 (1995).
22. T. Albrecht, W.W. Li, J. Ulstrup, W. Haehnel and P. Hildebrandt, *ChemPhysChem,* **6**, 961 (2005).
23. L. Riauba, G. Niaura, O. Eicher-Lorka and E. Butkus, *J. Phys. Chem. A*, **110**, 13394 (2006).
24. A. Kudelski and W. Hill, *Langmuir*, **15**, 3162 (1999).
25. S. K. Nandy and G. S. Kastha, *Indian J. Phys.*, **47**, 763 (1973).
26. H. Fleischer, Y. Dienes, B. Mathiasch, V. Schmitt and D. Schollmeyer, *Inorg. Chem.,* **44**, 8087 (2005).
27. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, *Gaussian 03 revision D01*, Gaussian Inc., Wallingford CT (2004).

Laurynas Riauba

SOLVATUOTO CISTEAMINO RAMANO SPEKTRŲ MODELIAVIMO TYRIMAI

S a n t r a u k a

Tyrinėtos solvatuotų molekulių Ramano spektrų skaičiavimo ypatybės. Modeline sistema buvo pasirinkti cisteamino hidrochlorido vandeniniai tirpalai. Atliekant skaičiavimus buvo naudojama Integralinių lygčių formalizmo poliarizuojamojo kontinuumo modelio (IEFPCM) metodika, o į struktūrų modelius įtrauktos papildomos vandens molekulės. Tokio tipo struktūrų geometrijos optimizacija gali trukti ilgą laiką dėl gana plokščio potencinės energijos paviršiaus. Analizuojant Ramano spektro pokyčius optimizacijos metu buvo nustatyta, kad gerokai mažiau optimizacijos žingsnių reikia norint gauti vidutiniškai tikslius rezultatus. Tyrinėtuose modeliuose po 10–15 optimizacijos žingsnių spektrinių parametrų skaičiavimo klaidos buvo ne didesnės kaip 5–10%. Taip pat buvo nustatyta, kad skaičiavimo klaidos gerai koreliuoja su vidutinėmis jėgomis, veikiančiomis atomus, ir su vidutiniais atomų padėties pokyčiais.