Computation of the amide I band of polypeptides and proteins using a partial Hessian approach

Nicholas A. Besley, and Katie A. Metcalf

Citation: The Journal of Chemical Physics **126**, 035101 (2007); doi: 10.1063/1.2426344
View online: https://doi.org/10.1063/1.2426344
View Table of Contents: http://aip.scitation.org/toc/jcp/126/3
Published by the American Institute of Physics

Computation of the amide I band of polypeptides and proteins using a partial Hessian approach

Nicholas A. Besley[a] and Katie A. Metcalf
School of Chemistry, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom

(Received 3 November 2006; accepted 1 December 2006; published online 17 January 2007)

A partial Hessian approximation for the computation of the amide I band of polypeptides and proteins is introduced. This approximation exploits the nature of the amide I band, which is largely localized on the carbonyl groups of the backbone amide residues. For a set of model peptides, harmonic frequencies computed from the Hessian comprising only derivatives of the energy with respect to the displacement of the carbon, oxygen, and nitrogen atoms of the backbone amide groups introduce mean absolute errors of 15 and 10 cm^{-1} from the full Hessian values at the Hartree-Fock/STO-3G and density functional theory EDF1/6-31G* levels of theory, respectively. Limiting the partial Hessian to include only derivatives with respect to the displacement of the backbone carbon and oxygen atoms yields corresponding errors of 24 and 22 cm^{-1}. Both approximations reproduce the full Hessian band profiles well with only a small shift to lower wave number. Computationally, the partial Hessian approximation is used in the solution of the coupled perturbed Hartree-Fock/Kohn-Sham equations and the evaluation of the second derivatives of the electron repulsion integrals. The resulting computational savings are substantial and grow with the size of the polypeptide. At the HF/STO-3G level, the partial Hessian calculation for a polypeptide comprising five tryptophan residues takes approximately 10%–15% of the time for the full Hessian calculation. Using the partial Hessian method, the amide I bands of the constituent secondary structure elements of the protein agitoxin 2 (PDB code 1AGT) are calculated, and the amide I band of the full protein estimated. © *2007 American Institute of Physics*. [DOI: 10.1063/1.2426344]

INTRODUCTION

Determining the structure of polypeptides and proteins represents a fundamental challenge in biophysical chemistry. X-ray crystallography can provide structural resolution at an atomic level. However, this technique is often unsuitable since many proteins cannot be crystallized. Furthermore, X-ray crystallography provides a static picture of the structure. In contrast, time-resolved spectroscopic probes of protein structure provide dynamic structural information and can be applied to the study of the evolution of protein structure. This can potentially provide insight into the mechanism of protein folding.[1–3] Nuclear magnetic resonance spectroscopy provides information at the atomic level but cannot yet be measured on a submicrosecond time scale. Consequently, there is continued interest in low-resolution spectroscopic probes of protein structure. Infrared (IR) spectroscopy is well established as a probe of protein secondary structure and can be measured on the picosecond time scale.[4–7] The amide I band is the most useful for structure determination and arises predominantly from the C=O stretch of the amide carbonyl groups. In general, α-helices exhibit a peak at approximately 1650 cm^{-1}. Antiparallel β-sheets are characterized by a peak in the range of 1620–1640 cm^{-1}.[8] There is also interest in the IR spectroscopy of smaller polypeptides. Recently, two-dimensional IR spectroscopy has been used to study polypeptides[9] and Fourier transform IR spectroscopy has been used to probe the secondary structure of polypeptides.[10,11]

In conjunction with experimental studies, there is an extensive effort toward theoretical modeling of the amide bands, since this is crucial for the interpretation and understanding of experimental spectra. Historically, analysis of protein IR spectra has been based on empirical relationships between the amide I band profile and protein structure.[12–15] Recently, there has been considerable effort to develop quantitative prediction of the amide I band from protein structure. Vibrational spectra can be computed routinely using *ab initio* or density functional theory (DFT).[16] The vast majority of work in this area considers harmonic frequencies that are compared directly to experiment, although the computation of anharmonic frequencies is becoming increasingly feasible.[17] However, even the calculation of harmonic frequencies is not extended readily to polypeptides or proteins due to the computational cost of computing and diagonalizing the Hessian matrix. Cho and co-workers reported restricted Hartree-Fock (HF) calculations with the 6-311++G** basis set for a number of small polypeptides,[18] while much larger polypeptides can be studied using semiempirical quantum chemical methods.[19] However, recent DFT calculations of the vibrational frequencies of the pentapeptide [Leu]enkephalin represent the current limit of quantitatively accurate *ab initio* calculations.[20]

Currently, the calculation of the amide I band of proteins is achieved through the transition dipole coupling (TDC)

[a]Electronic mail: nick.besley@nottingham.ac.uk

method.[21,22] In this approach, the C=O stretch is treated as an oscillator with the coupling to the other vibrational modes neglected. A Hamiltonian matrix is constructed where the diagonal elements correspond to the frequencies of the individual amide oscillators and the off-diagonal elements describe the coupling between the different oscillators. Diagonalization of this matrix gives the coupled amide I frequencies of the protein. This method requires parameters describing the amide I frequencies of the peptide units and their coupling. These parameters can be obtained from quantum chemical calculations of model amides, such as *N*-methylacetamide. Subsequently, the transition charge coupling (TCC) model was proposed which includes higher-order multipole moments.[23,24] These approaches fail to describe the short range through bond interactions. Recent work has addressed this problem by incorporating nearest neighbor coupling maps to account for these interactions within the TDC and TCC models.[25] These models are used widely, and many groups have developed parameters and studied the amide I band of polypeptides and proteins.[18,20,25–37] These calculations are becoming increasingly accurate and for a number of proteins provide qualitatively correct predictions of the amide I band profiles.

Despite the success of these approaches, it remains desirable to compute the IR spectroscopy of proteins directly using *ab initio* methods. Through such calculations, the individual environment of amide groups is described in detail providing, in principle, quantitatively accurate spectral band profiles. Furthermore, through analysis of the normal modes of such full *ab initio* calculations, the amide I band can be resolved with respect to the different secondary structure elements yielding a direct probe of the relationship between the spectral band and the underlying secondary structure. The direct calculation of the amide I band of proteins represents a formidable challenge to quantum chemistry. Proteins consist of hundreds or thousands of atoms and often the role of solvent is important. However, with the availability of supercomputers and the parallelization of quantum chemical algorithms, it is not unfeasible that such calculations may be realized in the future. In this paper, we report our recent work toward this aim. The amide I band has the characteristic of being localized on the carbonyl group. It is shown that computing a partial Hessian comprising second derivatives associated with the motion of the carbon, oxygen, and nitrogen atoms of the backbone carbonyl groups leads to large computational savings with the introduction of a relatively small error. Partial Hessian approaches have been used previously to study the spectroscopy of molecules adsorbed on surfaces and the localization of vibrational energy and entropy.[38–42] Through this approach the amide I band of much larger polypeptides can be studied from first principles with quantitative quantum chemical methods. To illustrate the type of system that can be studied, the amide I band arising from the secondary structure elements drawn from the protein agitoxin 2 (PDB code 1AGT) (Ref. 43) is computed.

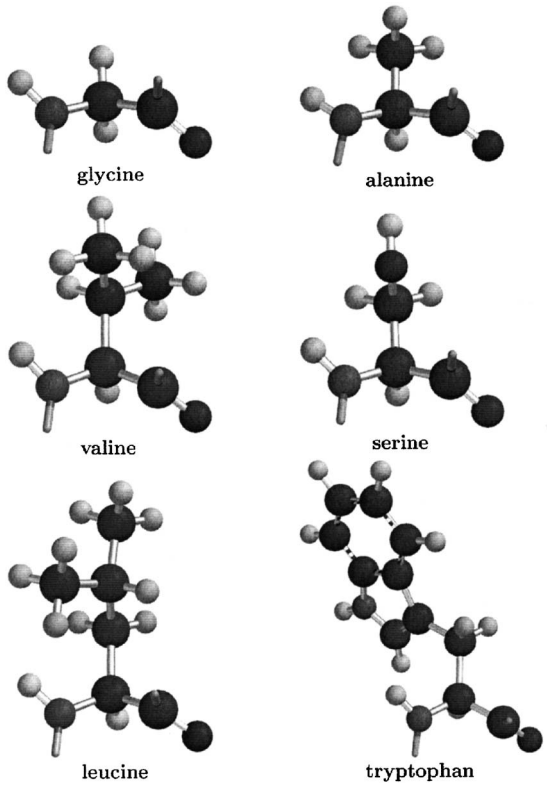

FIG. 1. Amino acids used in the test set of polypeptides.

COMPUTATIONAL DETAILS

Within the harmonic approximation, vibrational modes and frequencies are evaluated through computation and diagonalization of the Hessian matrix. The Hessian can be computed efficiently via the coupled perturbed Hartree-Fock (CPHF) equations[44] using the iterative procedure introduced by Pople *et al.*[45] The full Hessian has dimensions of $3N \times 3N$, where N is the number of atoms in the system. In this paper two different schemes are introduced; in the first a partial Hessian comprising the derivatives of the energy with respect to the movement of the carbon and oxygen atoms of the backbone carbonyl groups is computed. In the second scheme, the amide nitrogen atoms are also included. This results in Hessians with dimensions of $6N_{\mathrm{res}} \times 6N_{\mathrm{res}}$ and $9N_{\mathrm{res}} \times 9N_{\mathrm{res}}$, respectively, where N_{res} is the number of backbone residues. In the analytical computation of the Hessian, most computational work is involved in determining the derivatives of the density matrix and in the evaluation of the second derivatives of the electron repulsion integrals.[45] In our current implementation, the partial Hessian approximation is exploited in these two steps. This has been implemented within a developmental version of the Q-CHEM program package.[46]

The error introduced by the partial Hessian approximation is assessed through calculations on a set of glycine, alanine, valine, serine, leucine, and tryptophan polypeptides containing two to five peptide units (see Fig. 1). Following optimization of the structures, harmonic frequencies are computed. The calculations have been performed using restricted HF theory with the STO-3G basis set and DFT with the EDF1 exchange-correlation functional[47] and 6-31G* ba-

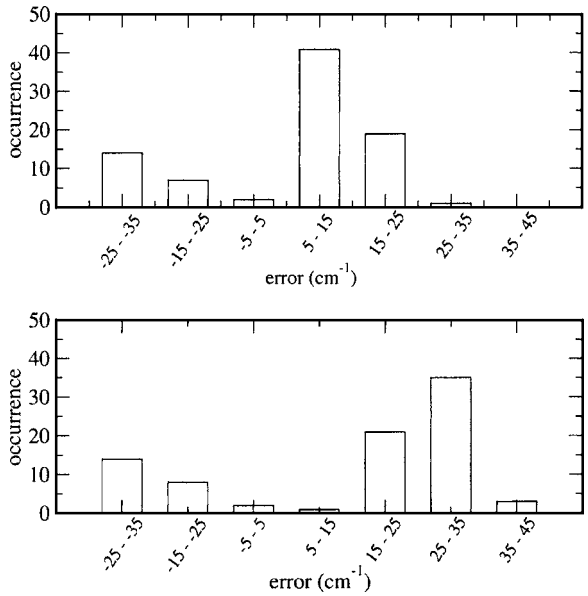

FIG. 2. Distribution of errors of the HF/STO-3G partial Hessian calculations. Lower panel: backbone carbon and oxygen atoms only. Upper panel: backbone carbon, oxygen, and nitrogen atoms.

sis set. Amide harmonic frequencies computed using EDF1/6-31+G* have been shown to agree closely with experiment.[48] EDF1/6-31G* also agrees well with experiment and avoids the problem of linear dependence in the basis set due to large numbers of diffuse basis functions. Due to the cost of the calculations, the four and five residue tryptophan peptides were not studied with the EDF1/6-31G* methodology. Computational timings are presented for the HF calculations and correspond to running on a single Intel 2.8 GHz processor with 2 Gbytes of memory. The protein agitoxin 2 contains α-helix, β-sheet, and turn secondary structure elements. The amide I band arising from these secondary structure elements extracted from the protein structure, with open valencies capped using with hydrogen, has been computed at the HF/6-31G level. These elements contain 10, 12, and 12 residues for the turn, helix, and β-sheet, respectively.

RESULTS AND DISCUSSION

Figure 2 shows the distribution of errors introduced by the partial Hessian approximation at the HF/STO-3G level of theory. Restricting the Hessian to include the derivatives of the amide carbon and oxygen atoms leads to a clustering of errors between 15–35 and −15 to −25 cm^{-1} with the largest error +41.8 cm^{-1}. Inclusion of the amide nitrogen atoms within the partial Hessian results in a significant improvement in accuracy. In particular, there is a decrease in large positive errors. The most frequent error lies within the range of 5–15 cm^{-1} and the largest error observed is +27.6 cm^{-1}. This improvement is reflected in the mean average deviation (MAD) from the full Hessian values shown in Table I. For partial Hessian restricted to the carbon and oxygen atoms, the MAD is 23.5 cm^{-1} and this improves to 14.5 cm^{-1} with the addition of the nitrogen atoms. The size of these errors is small in the context of other errors inherent in the calculations. The anharmonicity of the amide I mode is of the order

TABLE I. Mean absolute deviations (MADs) in the frequencies (in cm^{-1}) and intensities (in km mol^{-1}) in parentheses of the partial Hessian calculations from the full Hessian values.

Method	MAD (C,O)	MAD (C,O,N)
HF/STO-3G	23.5(34.3)	14.5(37.6)
EDF1/6-31G*	21.8(46.1)	9.6(40.5)

of 10–18 cm^{-1}. Consequently, the magnitude of the error of the partial Hessian approximation is similar to the error arising from the harmonic approximation.

The calculated intensities of the amide I modes are more sensitive to the partial Hessian approximation than the frequencies. At the HF/STO-3G level, the MADs in the computed intensities are 34.3 and 37.6 km mol^{-1} for the carbon/oxygen and carbon/oxygen/nitrogen Hessian approximations, respectively. Surprisingly, the calculation with the larger Hessian gives worse agreement with the full Hessian values. However, the accuracy of the partial Hessian approximation is most easily assessed through comparison of the computed amide I band profiles, since the band profiles are used in structural assignments. These are generated by representing each vibrational mode with a Gaussian with a full width half maximum of 10 cm^{-1}. The HF/STO-3G computed band profiles for the largest of each of the polypeptides are shown in Fig. 3. The amide I bands are shifted to lower wave number, but the general shape of the full Hessian bands is reproduced by both partial Hessian approximations. In all cases the larger partial Hessian leads to band profiles that are significantly closer to the full Hessian bands. The exception is pentaglycine, for which the partial Hessian bands differ significantly from the full Hessian. This difference arises from the computed intensity of one of the modes contributing to the lower energy band, which is greater in the partial Hessian

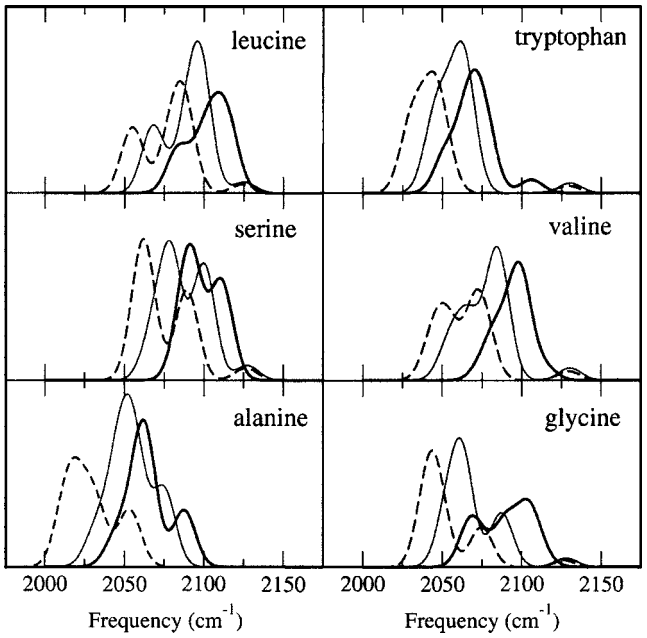

FIG. 3. HF/STO-3G amide I band profiles for the pentapeptides. Bold line—full Hessian, solid line—backbone carbon/oxygen/nitrogen Hessian, and broken line—backbone carbon/oxygen Hessian.

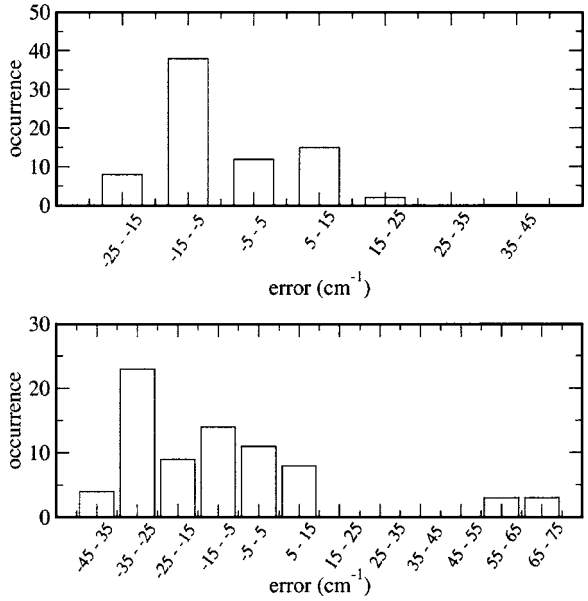

FIG. 4. Distribution of errors of the EDF1/6-31G* partial Hessian calculations. Lower panel: backbone carbon and oxygen atoms only. Upper panel: backbone carbon, oxygen, and nitrogen atoms.

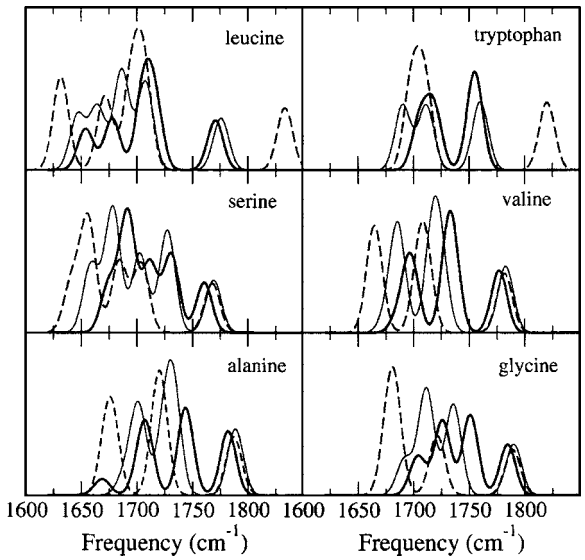

FIG. 5. EDF1/6-31G* amide I band profiles for the pentapeptides and tryptophan tripeptide. Bold line—full Hessian, solid line—backbone carbon/oxygen/nitrogen Hessian, and broken line—backbone carbon/oxygen Hessian.

calculations. This case appears atypical, and for the smaller glycine peptides the full Hessian band profiles are well reproduced. The partial Hessian could be expanded to reduce the error further. The amide I mode also has some contribution from the amide hydrogen. Our calculations indicate that inclusion of the amide hydrogen leads to a small improvement in the computed frequencies of intensities. Enlargement of the Hessian will increase the cost of the calculations, and the additional accuracy obtained from including amide hydrogen probably does not justify this increase.

At the EDF1/6-31G* level of theory, the MADs of the partial Hessian calculations compared to the full Hessian are 21.8 and 9.6 cm^{-1} for the carbon/oxygen and carbon/oxygen/nitrogen partial Hessian calculations, respectively. These are lower than the corresponding errors using HF/STO-3G, suggesting a weaker coupling to the other vibrational modes. The distribution of errors observed for the EDF1/6-31G* calculations is shown in Fig. 4. For the Hessian restricted to the carbonyl carbon and oxygen atoms there are a number of large positive errors, with the largest observed error of +67.81 cm^{-1}. These modes with a very large error arise from the C═O stretch of carbonyl groups on the end residue of the larger leucine and tryptophan polypeptides. These carbonyl groups are not involved in hydrogen bonding and there is no clear structural reason why they are more sensitive to the partial Hessian approximation. These outliers are no longer present with the enlarged partial Hessian, and for this Hessian there is a narrow distribution of errors, with the largest error reduced to +24.41 cm^{-1}. The MADs of the computed intensities are 46.1 and 40.5 km mol^{-1} for the carbon/oxygen and carbon/oxygen/nitrogen partial Hessian approximations. The computed band profiles are shown in Fig. 5. The computed frequencies of the amide bands lie in the range of 1650–1800 cm^{-1}. This is in reasonable agreement with the values from experiment and demonstrates the increased accuracy of the better level of theory. The band profiles arising from the partial Hessian that includes the backbone nitrogen atoms are in good agreement with the full Hessian calculations. The band shapes are predicted well with a small shift. For the carbon/oxygen partial Hessian, the computed band profiles are significantly worse and poorer than similar calculations using HF/STO-3G. In particular, for leucine and tryptophan spurious bands at high wave numbers are evident.

The benefit of the partial Hessian approximation is the substantial reduction in the computational cost in terms of the time and memory required. This is illustrated in Fig. 6, which shows the time for the computation and diagonalization of the partial Hessian as a percentage of the time for the full Hessian calculation. This is shown for the HF/STO-3G calculations. For the polypeptides with small side chains, for example, glycine, there is a reasonably modest saving of about 20%. However, for polypeptides with large side chains, the savings increase to over 80%. For the tryptophan

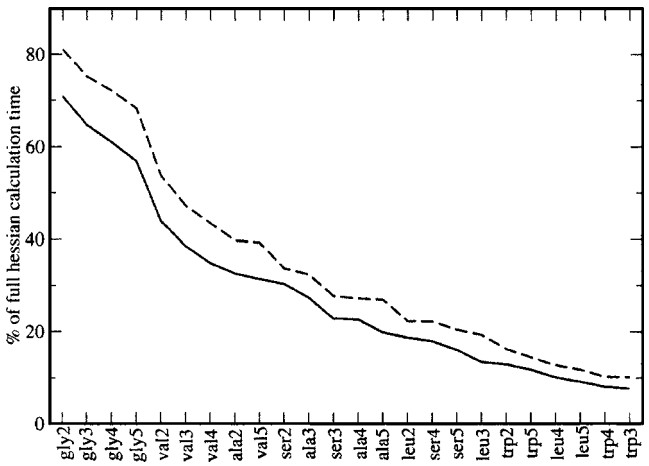

FIG. 6. Computational savings of the partial Hessian approach. Solid line—backbone carbon/oxygen Hessian and broken line—backbone carbon/oxygen/nitrogen Hessian.

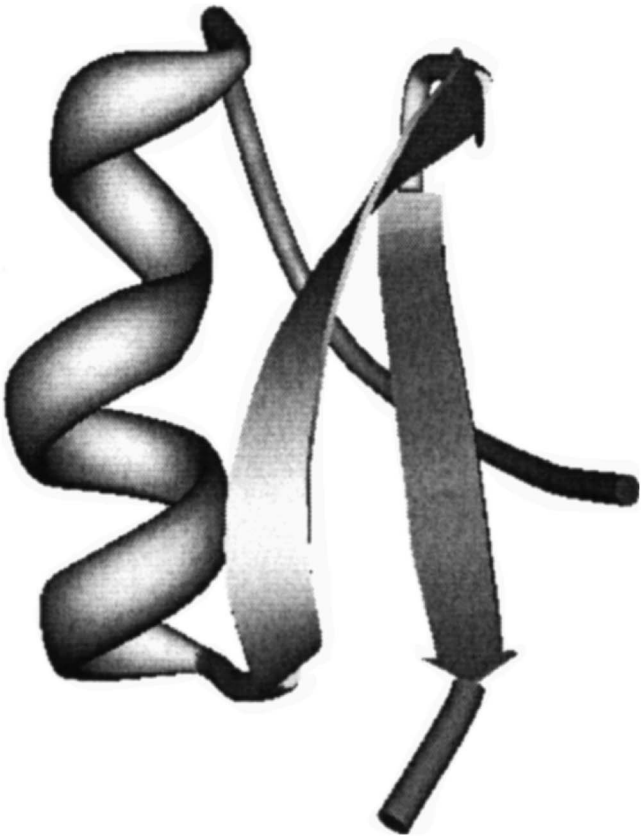

FIG. 7. Agitoxin 2 (PDB code 1AGT).

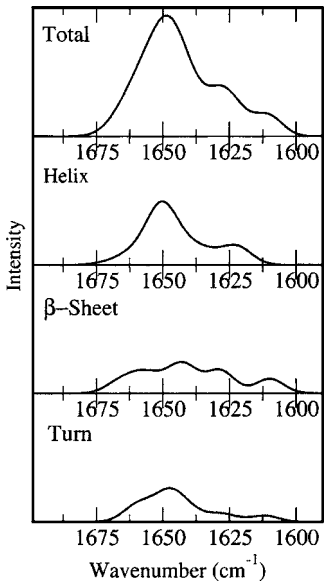

FIG. 8. Partial Hessian HF/6-31G amide I band profiles for the secondary structure elements of 1AGT. The total band profile is represented as a sum of the amide I bands for the individual secondary structure elements.

pentapeptide, the time for the full Hessian calculation is 23 136 s compared to 2731 and 3358 s for the carbon/oxygen and carbon/oxygen/nitrogen partial Hessian calculations, respectively. For this system, there is approximately an 84% saving in the time for the solution of the CPHF equations and the number of electron repulsion integrals is reduced by about a factor of 10. In general, the savings increase with the length of the polypeptide. There is some deviation from this for the tryptophan peptides. This arises from differences in the speed of the convergence of the iterative coupled perturbed self-consistent field procedure. The smaller partial Hessian calculations led to increased computational savings. However, for the carbon/oxygen/nitrogen partial Hessian approximation the savings are also substantial. Consequently, the partial Hessian approach avoids the rapid increase in computational cost of standard harmonic frequency calculations, allowing larger polypeptides to be studied more readily.

The protein agitoxin 2 (1AGT), shown in Fig. 7, has three distinct secondary structure elements, an α-helix, β-sheet, and turn. The amide I band profiles for these secondary structure elements computed with the partial Hessian approximation comprising the backbone carbon, oxygen, and nitrogen atoms within HF/6-31G theory are shown in Fig. 8. The largest fragment is the β-sheet, which contains 190 atoms. The spectra have been shifted by -135 cm^{-1} to give spectra that are in agreement with the characteristic amide I region of proteins. Harmonic frequencies computed with HF theory are well known to be too high compared with experiment. Furthermore, the calculations do not take into account solvent, which will also affect the computed amide I frequencies. The amide I band for the α-helix fragment consists of a broad band at 1650 cm^{-1} with a shoulder at 1623 cm^{-1}. This band is characteristic of α-helices and was also evident in TDC calculations of the mainly α-helical protein myoglobin by Choi *et al.*[28] TDC calculations were also reported for flaxodoxin, which contains α-helix and β-sheet secondary structure. The calculations predicted three bands at 1653, 1636, and 1628 cm^{-1} with the band at lower wave number localized on the carbonyl groups of the β-sheets. The partial Hessian calculations for the β-sheet shows a fairly broad feature, but a distinct band at about 1610 cm^{-1} can be distinguished. Bands in the 1610^{-1}–1630 cm^{-1} region are characteristic of β-sheet containing proteins. The spectrum for the turn is also fairly broad with significant intensity in the 1650 cm^{-1} region. Analysis of the normal modes from the partial Hessian calculation can provide an approximate picture of the nature of the vibrational modes. The intense band of the α-helix comprises an in phase vibration delocalized over a number of carbonyl groups. In contrast, the band at low wave number band in the β-sheet spectrum is localized largely on one carbonyl group in the loop region of the β-sheet.

In order to estimate the amide I band for the protein, we have summed the spectra for the three secondary structure elements. This assumes that the coupling between the amide I modes of the individual secondary structure elements is negligible. The resulting spectrum has three bands and is in good agreement with previous work on mixed α/β proteins.[28] The largest band at 1650 cm^{-1} has contributions from all secondary structure elements, but the largest contribution is clearly from the α-helix. The band at 1630 cm^{-1} has roughly equal contributions from the α-helix and β-sheet, while the band at 1610 cm^{-1} arises from the β-sheet. The time to compute the frequencies for the β-sheet fragment was approximately 36 h. While this time can be

reduced through parallelization, computational time may make the study of multiple structures drawn from a molecular dynamics simulation prohibitive. However, the partial Hessian approach may facilitate the refinement of the computationally less demanding TDC and TCC approaches.

CONCLUSIONS

The amide I band is the most useful in protein structure determination and arises from carbonyl stretching modes that are localized on the carbonyl group. This property has been exploited to reduce the computational cost of the *ab initio* calculation of the amide I band by constructing and diagonalizing a partial Hessian comprising derivatives associated with the atoms of the carbonyl group. Restricting the Hessian to include carbon, oxygen, and nitrogen atoms leads to MADs of only 14.5 and 9.6 cm^{-1} from the full Hessian values at the HF/STO-3G and EDF1/6-31G* levels of theory, respectively. These errors are small, particularly in the context of other errors inherent in the calculations. The computed intensities are more sensitive to the partial Hessian approximation, but the amide I band profiles are reproduced well. The more severe partial Hessian approximation that restricts the Hessian to the amide carbon and oxygen atoms is less reliable. The MAD increases to 23.5 and 21.8 cm^{-1}. The corresponding band profiles show more deviation from the full Hessian bands but in many cases provide a reasonable description of the amide I band profile.

The amide I bands of the secondary structure elements of the protein agitoxin 2 have been computed within the partial Hessian approximation using HF/6-31G. The amide I band for the protein is estimated as a sum of these bands. The shifted spectra are in good agreement with previous work using the TDC method.[28] The amide I band for the full protein has three distinct bands at 1650, 1630, and 1610 cm^{-1}. The band at 1650 cm^{-1} arises predominantly from the α-helix but has significant contributions from the β-sheet and turn. The band at 1630 cm^{-1} has similar contributions from the α-helix and β-sheet, while the band at 1610 cm^{-1} arises from the β-sheet and is characteristic of β-sheet containing proteins.

The partial Hessian approach to computing the amide I band provides a method that can extend the direct *ab initio* calculation of amide I bands to larger polypeptides. The current limitation of the code is the memory required for the calculations. All calculations presented here use the serial version of the Q-CHEM software. However, analytical derivatives have been implemented within a parallel version of the code.[49] Incorporation of the partial Hessian approach within parallel Q-CHEM will further extend the size of polypeptide that can be studied. This will allow more accurate levels of theory to be used so it may be no longer necessary to shift the spectrum to get agreement with experiment. Furthermore, this will make the *ab initio* calculation of the amide I bands of the complete protein a realistic possibility. This will make the assumption that there is negligible coupling between the amide I modes of the different secondary structure elements, used in the calculations of 1AGT presented here, unnecessary. These calculations will provide quantitative insight into the nature of the amide I band and may also prove useful in refining TDC and TCC approaches. This work is currently underway.

ACKNOWLEDGMENTS

One of the authors (N.A.B.) is grateful to the Engineering and Physical Sciences Research Council for funding, in particular, the award of an Advanced Research Fellowship (GR/R77636), and to the University of Nottingham for access to its high performance computing facility. The authors thank Professor J. D. Hirst and Dr. T. M. Watson for useful discussions.

[1] F. M. Hughson, P. E. Wright, and R. L. Baldwin, Science **29**, 1544 (1990).
[2] Y. V. Griko and P. L. Privalov, J. Mol. Biol. **235**, 1318 (1994).
[3] P. A. Jennings and P. E. Wright, Science **262**, 892 (1993).
[4] H.-U. Gremlich and B. Yan, *Infrared and Raman Spectroscopy of Biological Materials* (Dekker, New York, 2000).
[5] W. K. Surewicz and H. H. Mantsch, in *Spectroscopic Methods for Determining Protein Structure in Solution*, edited by H. A. Havel (VCH, New York, 1996), p. 135.
[6] Z. Chi and S. A. Asher, Biochemistry **37**, 2865 (1998).
[7] Z. Chi, X. G. Chen, J. S. W. Holtz, and S. A. Asher, Biochemistry **37**, 2855 (1998).
[8] M. Jackson, P. I. Haris, and D. Chapman, J. Mol. Struct. **214**, 329 (1989).
[9] J. Wang, J. Chen, and R. M. Hochstrasser, J. Phys. Chem. B **110**, 7545 (2006).
[10] R. Schweitzer, Biophys. J. **83**, 523 (2002).
[11] R. Schweitzer, J. Phys. Chem. B **108**, 16965 (2004).
[12] A. Elliot and B. J. Ambrose, Nature (London) **165**, 921 (1950).
[13] T. Miyazawa and E. R. Blout, J. Am. Chem. Soc. **83**, 712 (1961).
[14] S. Krimm and J. Bandekar, Adv. Protein Chem. **38**, 181 (1986).
[15] F. Dousseau and M. Pezolet, Biochemistry **29**, 8771 (1990).
[16] F. Jensen, *Introduction to Computational Chemistry* (Wiley, Chichester, 1999).
[17] Y. Wang and R. M. Hochstrasser, J. Phys. Chem. B **110**, 3798 (2006).
[18] J. H. Choi, S. Y. Ham, and M. Cho, J. Phys. Chem. B **107**, 9132 (2003).
[19] J.-H. Choi, S. Hahn, and M. Cho, Int. J. Quantum Chem. **104**, 616 (2005).
[20] T. M. Watson and J. D. Hirst, Phys. Chem. Chem. Phys. **6**, 2580 (2004).
[21] S. Krimm and Y. Abe, Proc. Natl. Acad. Sci. U.S.A. **69**, 2788 (1972).
[22] H. Torii and M. Tasumi, J. Chem. Phys. **96**, 3379 (1992).
[23] P. Hamm and S. Woutersen, Bull. Chem. Soc. Jpn. **38**, 181 (2002).
[24] M. Swart, P. T. van Duijnen, and J. G. Snijders, J. Comput. Chem. **22**, 79 (2001).
[25] T. l. C. Jansen, A. G. Dijkstra, T. M. Watson, J. D. Hirst, and J. Knoester, J. Chem. Phys. **125**, 044312 (2006).
[26] H. Torii and M. Tasumi, J. Chem. Phys. **97**, 86 (1992).
[27] D. A. Dixon, K. D. Dobbs, and J. J. Valentini, J. Phys. Chem. **98**, 13435 (1994).
[28] J. H. Choi, S. Ham, and M. Cho, J. Chem. Phys. **117**, 6821 (2002).
[29] S. Cha, S. Ham, and M. Cho, J. Chem. Phys. **117**, 740 (2002).
[30] P. Bour and T. A. Keiderling, J. Chem. Phys. **119**, 11253 (2003).
[31] S. Ham, J. H. Kim, H. Lee, and M. H. Cho, J. Chem. Phys. **118**, 3491 (2003).
[32] S. Ham, S. Cha, J. H. Choi, and M. Cho, J. Chem. Phys. **119**, 1451 (2003).
[33] J. R. Schmidt, S. A. Corcelli, and J. L. Skinner, J. Chem. Phys. **121**, 8887 (2004).
[34] T. M. Watson and J. D. Hirst, Mol. Phys. **103**, 1531 (2005).
[35] T. Hayashi, W. Zhuang, and S. Mukamel, J. Phys. Chem. A **109**, 9747 (2005).
[36] T. l. C. Jansen and J. Knoester, J. Chem. Phys. **124**, 044502 (2006).
[37] H. S. Chung and A. Tokmakoff, J. Phys. Chem. B **110**, 2888 (2006).
[38] S. Jin and J. D. Head, Surf. Sci. **318**, 204 (1994).
[39] M. D. Calvin, J. D. Head, and S. Jin, Surf. Sci. **345**, 161 (1996).
[40] J. D. Head, Int. J. Quantum Chem. **65**, 827 (1997).

[41] J. D. Head and Y. Shi, Int. J. Quantum Chem. **75**, 815 (1999).
[42] J. D. Head, Int. J. Quantum Chem. **77**, 350 (2000).
[43] A. M. Krezel, C. Kasibhatla, P. Hidalgo, R. Mackinnon, and G. Wagner, Protein Sci. **4**, 1478 (1996).
[44] J. Gerrat and I. M. Mills, J. Chem. Phys. **49**, 1719 (1968).
[45] J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. Quantum Chem., Quantum Chem. Symp. **13**, 225 (1979).
[46] Y. Shao, L. Fusti-Molnar, Y. Jung *et al.*, Phys. Chem. Chem. Phys. **8**, 3172 (2006).
[47] R. D. Adamson, P. M. W. Gill, and J. A. Pople, Chem. Phys. Lett. **284**, 6 (1998).
[48] T. M. Watson and J. D. Hirst, J. Phys. Chem. A **106**, 7858 (2002).
[49] P. P. Korambath, J. Kong, T. R. Furlani, and M. Head-Gordon, Mol. Phys. **100**, 1755 (2002).