Application of the Systematic Molecular Fragmentation by Annihilation Method to *ab Initio* NMR Chemical Shift Calculations

Rika Kobayashi,*,[†] Roger D. Amos,[†] David M. Reid,[‡] and Michael A. Collins[‡]

International Centre for Quantum and Molecular Structure, College of Sciences, Shanghai University, Shanghai 200444, P. R. China
[†]ANU Supercomputer Facility, Leonard Huxley Building 56, Mills Road, Canberra, ACT 2601, Australia
[‡]Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia

S *Supporting Information*

ABSTRACT: NMR is a powerful tool for obtaining information on the structural characterization and dynamics of proteins, and nucleic acids, and their complexes. The complexity of the spectra is such that elucidation through computational simulation is a much desired thing. However, the size of most structures of interest is such that they remain out of reach of accurate quantum chemical techniques. Fragmentation methods have been shown to be a viable means of reducing the cost of *ab initio* calculations to enable the prediction of molecular properties of large systems to chemical accuracy. We look at the systematic molecular fragmentation by annihilation method for a model peptide system and show that this procedure reproduces the shielding constants of a full calculation at only a fraction of the cost. Discussion of the considerations needed in applying this method is discussed and comparison made with the results of the similar fragment molecular orbital and ONIOM methods.
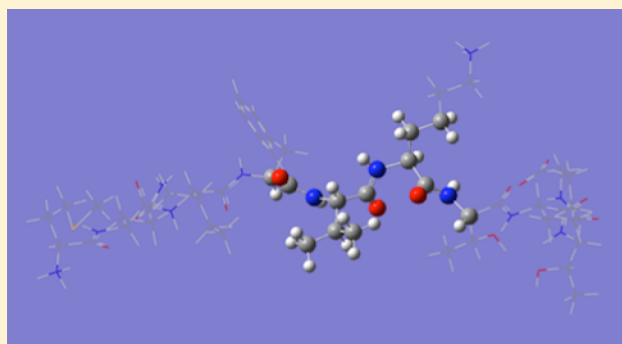
INTRODUCTION

Ab initio quantum chemistry is now firmly established as an invaluable tool to aid experimentalists, providing predictions with good accuracy for small molecules. Unfortunately, many of the systems of current interest are of a size that remains prohibitively expensive for accurate calculations, especially biomolecules and bulk systems. For this reason fragmentation methods are rapidly becoming a viable means of reducing the cost of calculations and enabling the prediction of molecular properties of large systems to chemical accuracy (for a review, see refs 1 and 2). These methods use the divide-and-conquer approach, where a chemical system is divided into smaller fragments, for which quantum chemistry calculations are performed on each, and the fragment properties then combined into a property for the whole. Further savings can be made as the fragments do not all need to be handled at the same level of theory, *e.g.* embedding methods and QM/MM. Such methods are particularly suited to the calculation of NMR spectra, as they depend on the "chemical locality" of macromolecular systems, which assumes that the local region of a macromolecule is only weakly influenced by the atoms that are far from the region of interest, NMR being a local property. There are a variety of fragment style methods that have been applied to NMR calculations, for example.[3−10] The present study is based on the application of the SMFA method to NMR by Reid and Collins.[11]

We recently presented an NMR study[12] on a model peptide system in which we showed that a customized fragment method based using some of the features of the fragment molecular orbital (FMO)[13] method, together with the locally dense basis approach, could reduce the cost of a calculation from 2165 h (on a 16 processor node of our Fujitsu CX400 x86_64 machine) to about 5 h (if all fragments were run simultaneously), while maintaining accuracy. In this follow-up paper we investigate an alternative fragmentation scheme, systematic molecular fragmentation analysis (SMFA),[14] to compare accuracy and cost. The two fragmentation schemes differ most significantly in the way the fragments are chosen and further investigations of this aspect naturally lent itself to an additional comparison with the ONIOM method.[15]

METHODOLOGY

The model system has been taken from the work of Gao et al.[10] and is the same as used in ref 12. It is built by taking the sequence of the first 10 residues of ubiquitin protein (MetGlnIlePhe-ValLysThrLeuThrGly), as shown schematically in Figure 1.

Gao et al. used a method based on the FMO method combined with gauge-including atomic orbitals, or continuous set of gauge transformations, to reproduce conventional ab initio NMR values for this system. We recently built on their study[12] by revisiting this system trying a different fragmentation pattern and applying the locally dense basis scheme[16] with Jensen's NMR shielding specialized pcSseg basis sets[17,18] which are
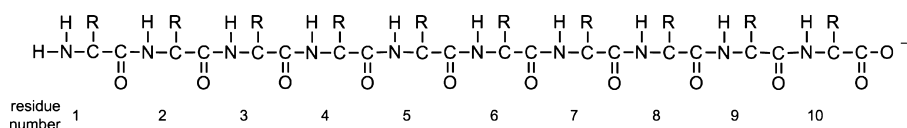
Figure 1. Scheme of the 10-residue polypeptide model system.

needed to eliminate the errors due to incomplete basis sets. In the same vein we investigate this strategy with SMFA and ONIOM.

Systematic Molecular Fragmentation by Annihilation (SMFA). The SMFA approximation has already been described in detail elsewhere (see ref 14. and references therein), so only a brief outline will be given. The NMR implementation with SMFA is given in ref 11. SMFA begins by considering a molecule as a set of functional groups connected by single bonds. The molecule, M, is then decomposed into N fragments F by systematically removing functional groups using different values of the parameter *Level*, as

$$M \rightarrow \sum_{i=1}^{N} c_i F_i \tag{1}$$

where F_i are overlapping fragments and c_i are integer fragment coefficients.

The value of *Level* reflects the proximity of eliminated groups, that preserves the bonding environment of each group to some extent. With increasing *Level* a more extensive bonding environment is retained. At *Level* = 1, the interaction of each group with its α substituents is included in the fragments. At *Level* = 2, α and β substituent interactions are included, and so on. The valency of the eliminated groups is maintained by appending hydrogen atoms along the original bond direction.

The total molecular electronic energy is given by

$$E = E_b + E_{nb} \tag{2}$$

where E_b denotes the "bonded energy" of fragment F_i, evaluated in the presence of embedded charges

$$E_b = \sum_{i=1}^{N} c_i E(F_i) \tag{3}$$

with $E(F_i)$ being the energy of the ith fragment, and E_{nb} denotes the "non-bonded energy."

The embedded charges represent the charge distribution of the atomic centers in the functional groups not contained in the fragment and evaluated by natural population analysis as described in Appendix B of ref 19.

The NMR shielding coefficients are obtained by taking the isotropic component of the chemical shielding tensor of a nucleus n:

$$\sigma_{\alpha\beta}^n = \frac{\partial^2 E}{\partial \mu_\alpha^n \partial B_\beta} \tag{4}$$

where B is an external magnetic field and μ^n is the nuclear magnetic moment, as

$$\sigma^n = \frac{1}{3} \sum_\alpha \sigma_{\alpha\alpha}^n \tag{5}$$

Nonbonded interactions due to electrostatic charges are absorbed into the fragment energies; i.e., the fragment is embedded in the field of the remote charges. The effect of long-range interactions due to magnetic fields on the chemical

shielding tensors of the fragment nuclei can be directly evaluated using perturbation theory, commonly known as "McConnell corrections":[20]

$$\delta\sigma^n = \frac{\chi_j}{r_{nj}^3} - \frac{\chi_j \cdot r_{nj} \cdot r_{nj}^T}{r_{nj}^5} \tag{6}$$

where χ_j is the average magnetic susceptibility tensor of fragment j and r_{nj} is the vector connecting atom n (in fragment i) with fragment j.

It was found in an earlier study on NMR within SMFA,[11] and confirmed in this one, that atoms involved in hydrogen bonding displayed lower errors when hydrogen bonding was included explicitly between groups for the purpose of fragmentation, rather than through the nonbonded correction. So regarding the nonbonded interactions, the inclusion of background multipoles up to second order using GDMA2[21] and the McConnell correction is adequate.

The total equation for the shielding of nucleus n then becomes

$$\sigma^n = \sum_i c_i \sigma^n(F_i) + \sum_j \frac{1}{3} \text{tr}\left(\frac{\chi_j}{r_{nj}^3} - \frac{\chi_j \cdot r_{nj} \cdot r_{nj}^T}{r_{nj}^5} \right) \tag{7}$$

where the sum over i in the first term is over all fragments which contain nucleus n, and the sum over j in the second term is over all fragments which do *not* contain atom n.

ONIOM. The ONIOM (order-N-in-order-M hybrid method) was developed by Morokuma and co-workers in 2000[15] and has been under continual refinement ever since. Though not strictly a fragmentation approach, it works on the same principles and provides an alternative scheme for comparison. Essentially, the ONIOM method partitions the system into 2 or 3 layers representing the region of interest handled with a high-accuracy method, a cheaper low-layer for environmental effects and optionally an intermediate layer to better describe neighboring interactions. Thus, ONIOM and QM/MM could be thought of as fragment models with only 2 or 3 fragments.

CALCULATIONS

The 10-peptide model system was input into the SMFA program and GIAO NMR calculations run with B3LYP using 6-31G*, pcSseg-1, and pcSseg-3[17,18] basis sets using the Gaussian program.[22] It was previously established that a basis set at least as large as pcSseg-3 is needed to eliminate errors due to basis incompleteness. Note that unlike FMO and ONIOM, the SMFA is an automated procedure and the fragmentation is determined by a systematic consideration of functional groups. There is provision for user-intervention in fragment choice but this was not done in this case. The automated procedure broke down the peptide system into:

49 fragments at *Level* = 0 (average size of fragments: 5 atoms)

95 fragments at *Level* = 1 (average size of fragments: 7 atoms)

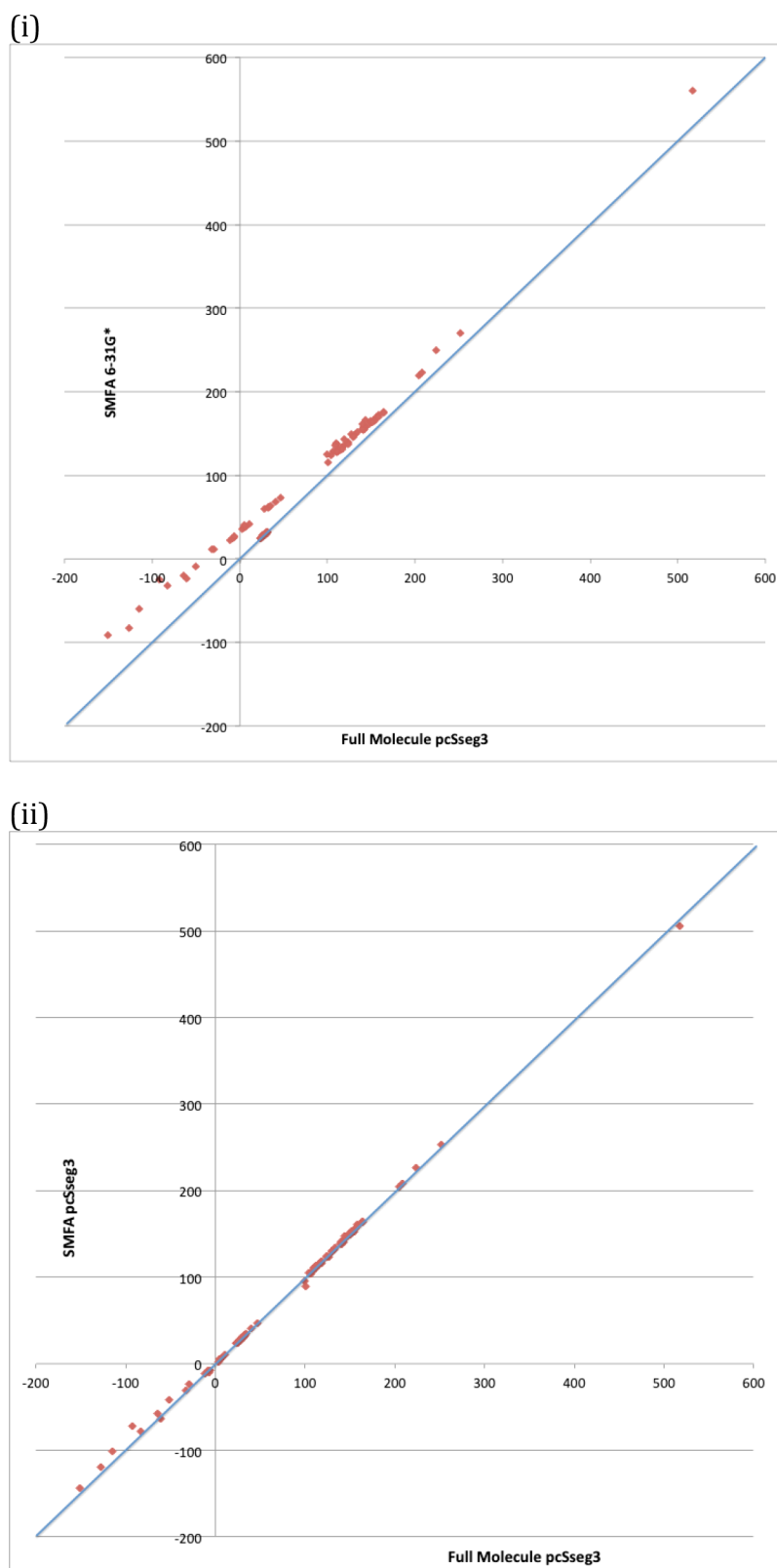79 fragments at *Level* = 3 (average size of fragments: 18 atoms)

Figure 2. (i) 6-31G(d) and (ii) pcSseg-3 isotropic shielding (in ppm) using the SMFA method compared to pcSseg-3 full molecule (=basis set limit).

This compares with the 8 fragments of average size 54 atoms in our previous study.[12]

Previous investigations[11] have found SMFA at *Level* = 3 fragmentation to be sufficient for the calculation of NMR shielding constants.

Figure 2 shows a plot of the SMFA shielding constants (absolute shieldings not chemical shifts) for all atoms in the molecule i.e. C,H,O,N and one sulfur atom, using B3LYP with (i) 6-31G(d) and (ii) pcSseg-3 basis sets versus the full-molecule pcSseg-3 calculation which we take to be the basis set

limit. Points on the line would agree exactly with the full molecule calculation. We have not used other functionals as the purpose of this study is to investigate the fragmentation algorithm, and as the errors due to basis set incompleteness are larger than those due to the limitations of the method chosen.

Figure 2i shows that the 6-31G(d) basis set consistently overestimates the values of the shielding constants confirming our belief from our earlier study that this level of basis is not sufficient for high-accuracy NMR prediction. Figure 2(ii) shows excellent correlation between the fragmentation and full-molecule results demonstrating that the SMFA method is capable of adequately representing the full system. The points in Figure 2ii with negative shieldings that lie slightly off the line are the oxygen atoms which are extremely sensitive to methodology. The oxygen nuclei have greater basis set dependence than the other nuclei and may not be as reliable. This could be a general feature of peptide systems in that the negatively charge oxygen, in close proximity to the positively charged nitrogen, is sensitive to small changes in the density. Additionally, the terminal oxygen nuclei showed excessive variability probably due to fluctuations in electron density and were removed from the statistics.

Comparison with the pcSseg-3 results from the FMO approach from our previous investigation is given in Table 1.

Table 1. Root Mean Square Error and Maximum Absolute Error of NMR Shieldings (in ppm) Computed with FMO and SMFA Fragmentation Compared to the Basis Set Limit

	FMO (pcSseg-3)[a]		SMFA (6-31G(d))		SMFA (pcSseg-3)	
nuclei	RMS	max error	RMS	max error	RMS	max error
C	0.76	2.68	22.0	35.2	0.87	2.17
H	0.20	0.99	1.03	2.00	0.26	0.68
N	4.31	11.8	20.9	28.6	3.82	11.7
O	5.93	12.9	46.0	66.8	7.32	14.5

[a]From ref 12.

Examining the MAE and maximum error values shows we are reaching high accuracy and improving on the FMO results with the SMFA fragmentation with pcSseg-3 basis at least for the carbon and hydrogen nuclei. The main disagreement is in the negative shielding of the oxygen atoms.

However, more relevant, for the rationale of fragmentation methods being cost saving, is an examination of the total and average timings for the methods as given in Table 2. The full table of timings for the individual fragments is given in Table S2 of the Supporting Information. The size of the fragments were deemed such as to be more effectively run on 8 processors and so

Table 2. Calculation of Walltime and the Average Time per Fragment on a 16 Processor x86_64 Node for the Fragments in FMO-Based vs SMFA-Based Fragmentation

fragmentation	no. of fragments	av no. of atoms	av no. of basis functions	total walltime (h)	average walltime (h)
FMO[a]	8	54	2737	134	16.7
SMFA (Lev 1)	95	7	315	6	0.1
SMFA (Lev 3)	79	18	901	135	1.7

[a]From ref 12.

the 16 processor equivalent is given in Table 2 for the purpose of comparison.

The timings in Table 2 show that the precision of the full molecule 2165 h calculation on 16 processors as shown in Figure 2ii could be maintained at a computation time of 135 h, if all the fragments were calculated sequentially. This compares with 134 h from the FMO approach in ref 12. However, as the fragment calculations are independent of each other, if run concurrently, the time to solution can be reduced, the time required for the largest fragment being about 8 h. Comparison with our previous study shows SMFA to give smaller errors at similar cost if the computations are run sequentially, improving to about 1/3 of the wall time if run in parallel, depending on the number of processors used.

LOCAL BASIS SET EFFECTS

We have shown in earlier studies[12,23] that the calculation cost can be further reduced by exploiting the chemical locality of the NMR shielding through the use of locally dense basis sets. In ref 23 partitioning schemes for locally dense basis sets in the context of NMR shielding calculations were explored. Atom-based and group-based divisions were examined, finding that a three-tier hierarchy of accuracy for basis sets on the group of interest, adjacent groups and rest of molecule, was sufficient to achieve the desired level of accuracy at a minimal computational expense. Note, however, this work was not carried out within a fragmentation scheme. Because the nature of the SMFA algorithm involves overlapping fragments, unlike some other fragmentation methods, the division of basis sets is not obviously define—nuclei of interest can appear in more than one fragment (at most 21 for the model peptide). To take into account local basis set effects rigorously within the SMFA approach, this three-tiered division of basis sets should be carried out within each fragment, but this mechanism could not be easily incorporated within the current program due to its complexity.

In the SMFA approach, atoms can sometimes occur in just a few fragments, perhaps only one, but frequently in a larger number of overlapping fragments, so the NMR shielding for that region could be estimated by averaging the shielding of the values from those fragments calculated at pcSseg-3 level with the remaining fragments at pcSseg-1. Figure 3 shows a plot of the shielding constants at B3LYP with only the fragments of the nuclei of interest being carried out at the higher basis set. The correlation statistics and reduced timings from this approach are given in Table 3. Of course, for all atoms, this would be effectively calculating all fragments at pcSseg-3. However, if a region of interest is defined then only the subset of relevant fragments need be calculated. If the region of interest for any nuclei is defined as the union of all fragments that nuclei appears in, the number of Level = 3 fragments, to be calculated with the larger basis, becomes 46 of average size 28 atoms.

COMPARISON WITH ONIOM

As Table 3 shows attempting a simple locally dense basis set variant of the SMFA method diminished the accuracy somewhat making its validity dubious. However, this division of accuracy between region of interest and elsewhere was reminiscent of the ONIOM (or indeed any QM/MM) method encouraging comparison with that. QM/MM approaches have been used to study NMR, see, e.g., ref 9, with subjective choice of QM region. It can be viewed that SMFA can be used to determine the QM region in an automated systematic manner based on
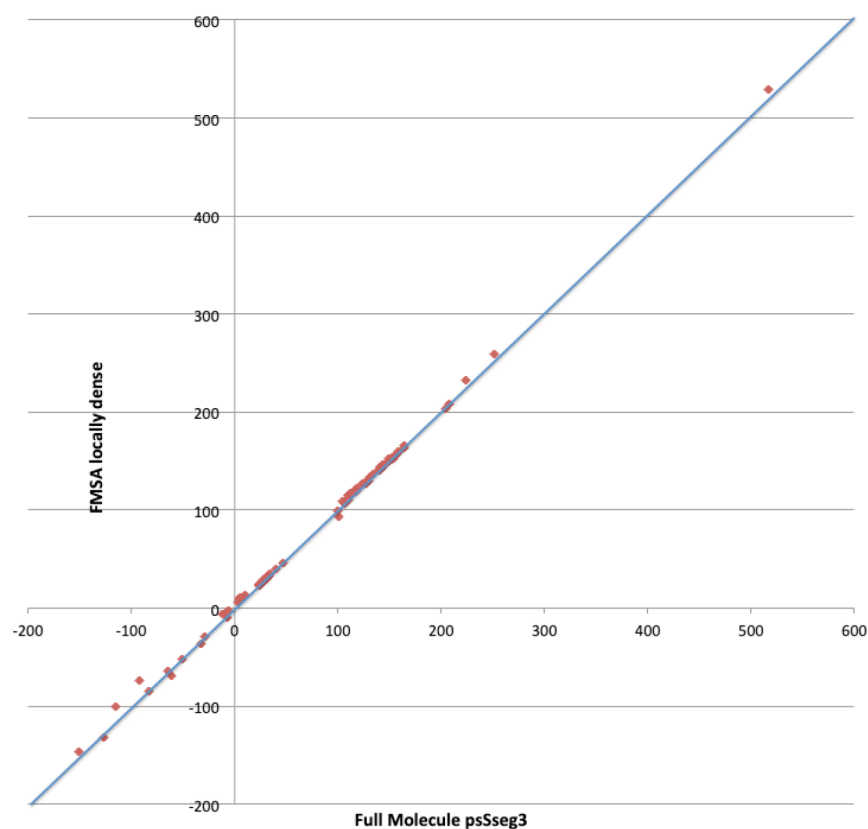
Figure 3. Comparison of locally dense SMFA shielding results (in ppm) compared to pcSseg-3 full molecule (=basis set limit).

Table 3. RMS Deviation and Maximum Error of Shieldings (in ppm) Using Locally Dense Basis Sets Compared to the Basis Set Limit

| nuclei | FMO (pcSseg-3)[a] | | | | SMFA (pcSseg-3) | | | |
| | original | | locally dense basis | | original | | locally dense basis | |
	RMS	max error	RMS	max error	RMS	max error	RMS	max error
C	0.76	2.68	0.72	2.72	0.87	2.17	2.85	5.83
H	0.20	0.99	0.28	1.02	0.26	0.68	0.29	0.78
N	4.31	11.8	4.72	12.1	3.82	11.7	2.79	7.10
O	5.93	12.9	5.71	11.2	7.32	14.5	8.16	18.25

[a]From ref 12.

chemical functionality. The "region of interest" for any nucleus in the system was decided as the union of all fragments containing it. Finding the union of all fragments containing any particular atom in the peptide led to 46 new superfragments covering the whole system. Figure 4 shows an example of the superfragment from the union of five Level 3 fragments determined by SMFA as contributing to the chemical shielding for atoms 78−94. This is actually the union of all the fragments which contain the central peptide unit in the Figure below. This fragment was put into the ONIOM routines in Gaussian 09, treated at "high-level" B3LYP/pcSseg-3 with the remainder of the molecule at "low-level" B3LYP/pcSseg-1. This differs from the SMFA method as follows:

(i) the low-level region of the molecule being treated as one fragment;

(ii) the treatment of linking atoms between fragments;

(iii) the charge model used for the environment;

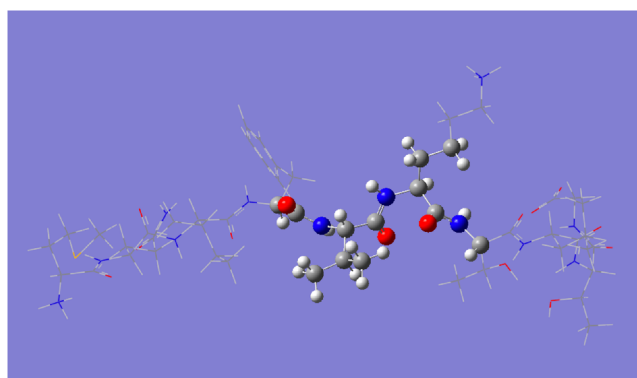(iv) the treatment of the interactions between fragments.

Figure 4. Superfragment from SMFA identifying all fragments contributing to a specified region of interest.

It should be clear that if one wants all the shieldings in the whole molecule then a fragment approach will be more economical than lots of mixed accuracy calculations. However,

there may be cases where one only wants to look at part of a system—for example looking at a molecule immersed in an explicit solvent, but only wanting the shieldings of the solute and not of the solvent. The SMFA technique may be of use in QM/MM or ONIOM more generally in that one may have a good idea of which are the most important atoms for the QM region while being unsure where the boundaries should be—in this case SMFA will help by attaching all the functionally important groups. In Table 4 are the shieldings for this fragment run as an

Table 4. Shieldings (in ppm) Using the SMFA Locally Dense Basis Approximation Compared to ONIOM for the Central Atoms in the Fragment Shown[a]

atom	full molecule[b]	fragment[b]	SMFA	SMFA-LD	ONIOM
78 (C)	111.24	110.89	112.01	116.15	113.08
79 (H)	26.07	27.94	27.95	28.21	27.85
80 (C)	141.31	141.28	141.83	143.96	140.47
81 (H)	29.59	29.56	29.69	29.71	29.47
82 (C)	152.57	152.83	153.61	153.38	152.90
83 (H)	31.16	31.00	31.36	31.11	31.19
84 (H)	30.44	30.18	30.41	30.36	30.50
85 (H)	29.73	30.72	29.98	29.75	29.70
86 (C)	158.03	159.13	159.68	160.34	160.05
87 (H)	30.88	30.82	31.01	30.70	30.82
88 (H)	30.77	30.48	31.00	30.55	30.86
89 (H)	30.51	30.52	30.60	30.42	30.48
90 (C)	5.67	2.99	5.86	10.47	5.51
91 (O)	−82.88	−77.73	−77.70	−84.80	−75.26
92 (N)	106.89	110.03	103.80	107.26	106.00
93 (H)	26.41	26.16	26.76	26.87	26.36
94 (C)	111.14	111.64	111.28	114.48	111.00

[a]The numbering corresponds to their position in the full molecule. [b]From ref 12.

ONIOM calculation, with pcSseg-3 on the QM region and pcSseg-1 everywhere else, compared to the other methods. The atom numbers are those in the full geometry, which is given in the Supporting Information, and correspond to the peptide unit in the fragment, and atoms in the chain immediately joined to it. The average statistics are in Table 5 and associated timings in Table 6. The full results using ONIOM are in the Supporting Information.

Table 5. RMS Deviation and Maximum Error of Shieldings (in ppm) Comparing the SMFA Locally Dense Basis Approximation to the ONIOM Approach

	SMFA LD		ONIOM	
	RMS	max error	RMS	max error
C	2.85	5.83	3.54	8.14
H	0.29	0.78	0.32	1.19
N	2.79	7.10	4.44	11.8
O	8.16	18.25	9.87	19.0

The results show the same pattern as for other methods—the carbon and hydrogen values are closely reproduced, but oxygen and nitrogen show some variability. Just looking at the values in Table 4, it is not obvious that the ONIOM method is a little less accurate than the other approaches, as comes out in the overall statistics (Table 5). It appears that ONIOM happens to do reasonably on this particular fragment, which was just chosen for aesthetic reasons. And again, Table 6 shows that for similar

Table 6. Calculation Walltime on 8 Processors for the ONIOM Fragment in Figure 4 vs the SMFA Fragments That Contribute to the QM Region

		no. of atoms	no. of basis functions	walltime (h)
ONIOM		36	2055	18.00
SMFA	frag 7	22	1116	5.06
	frag 15	19	978	3.23
	frag 18	19	978	2.77
	frag 19	19	978	2.79
	frag 22	19	978	3.10

accuracy the mixed-basis SMFA calculation can be run at much reduced computational wall time.

CONCLUSION

Two fragmentation schemes, the systematic molecular fragmentation by annihilation and the fragment molecular orbital method have been compared and contrasted for the calculation of NMR shielding constants. SMFA, which has the advantage of automatically choosing fragments based on chemical functionality, gives better agreement with full molecule calculations at much cheaper computational cost. However, the nature of the fragmentation scheme, producing overlapping fragments, is such that it can not exploit the use of locally dense basis sets in a straightforward manner. Nevertheless, using the fragmentation scheme to automatically determine a region of interest based on chemical functionality that could be handled at a high-level, analogous to QM/MM methods, compared favorably in accuracy and timing. We have shown that we are able to successfully simulate NMR calculations on large molecules at a fraction of the cost of conventional methods. The mechanism to do this—fragment choice, basis set choice, and inclusion of long-range interaction—has now been investigated, analyzed and set up, and we are now in a position to tackle real systems.

ASSOCIATED CONTENT

ⓢ Supporting Information
The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpca.8b09565.

> Full geometric coordinates of the system, timings for the fragment calculations, and all calculated shielding constants for each method discussed (PDF)

AUTHOR INFORMATION

Corresponding Author
*(R.K.) E-mail: Rika.Kobayashi@anu.edu.au.
ORCID ⓘ
Rika Kobayashi: 0000-0002-0672-833X
Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

■ REFERENCES

(1) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation Methods: A Route to Accurate Calculations on Large Systems. *Chem. Rev.* **2012**, *112*, 632−672.

(2) Collins, M. A.; Bettens, R. P. A. Energy-Based Molecular Fragmentation Methods. *Chem. Rev.* **2015**, *115*, 5607−5642.

(3) Jose, K. V. J.; Raghavachari, K. Fragment Based Approach for the Evaluation of NMR Chemical Shifts for Large Biomolecules Incorporating the Effects of the Solvent Environment. *J. Chem. Theory Comput.* **2017**, *13*, 1147−1158.

(4) Soss, S. E.; Flynn, P. J.; Iuliucci, R. J.; Young, R. P.; Mueller, L. J.; Hartman, J.; Beran, G. J. O.; Harper, J. K. Measuring and Modeling Highly Accurate 15-N Chemical Shift Tensors in a Peptide. *ChemPhysChem* **2017**, *18*, 2225−2232.

(5) Steinmann, C.; Bratholm, L. A.; Olsen, J. M. H.; Kongsted, J. Automated Fragmentation Polarizable Embedding Density Functional Theory (PE-DFT) Calculations of Nuclear Magnetic Resonance (NMR) Shielding Constants of Proteins with Application to Chemical Shift Predictions. *J. Chem. Theory Comput.* **2017**, *13*, 525−536.

(6) Zhu, T.; Zhang, J. Z. H.; He, X. Automated Fragmentation QM/MM Calculation of Amide Proton Chemical Shifts in Proteins with Explicit Solvent Model. *J. Chem. Theory Comput.* **2013**, *9*, 2104−2114.

(7) Zhao, D.; Song, R.; Li, W.; Ma, J.; Dong, H.; Li, S. Accurate Prediction of NMR Chemical Shifts in Macromolecular and Condensed-Phase Systems with the Generalized Energy-Based Fragmentation Method. *J. Chem. Theory Comput.* **2017**, *13*, 5231−5239.

(8) Swails, J.; Zhu, T.; He, X.; Case, D. A. AFNMR: Automated Fragmentation Quantum Mechanical Calculation of NMR Chemical Shifts for Biomolecules. *J. Biomol. NMR* **2015**, *63*, 125−139.

(9) He, X.; Wang, B.; Merz, K. M., Jr. Protein NMR Chemical Shift Calculations Based on Automated Fragmentation QM/MM Approach. *J. Phys. Chem. B* **2009**, *113*, 10380−10388.

(10) Gao, Q.; Yokojima, S.; Kohno, T.; Ishida, T.; Fedorov, D. G.; Kitaura, K.; Fujihira, M.; Nakamura, S. Ab Initio NMR Chemical Shift Calculations on Proteins Using Fragment Molecular Orbitals with Electrostatic Environment. *Chem. Phys. Lett.* **2007**, *445*, 331−339.

(11) Reid, D. M.; Collins, M. A. Calculating Nuclear Magnetic Resonance Shieldings Using Systematic Molecular Fragmentation by Annihilation. *Phys. Chem. Chem. Phys.* **2015**, *17*, 5314−5320.

(12) Amos, R. D.; Kobayashi, R. Ab Initio NMR Chemical Shift Calculations Using Fragment Molecular Orbitals and Locally Dense Basis Sets. *J. Phys. Chem. A* **2016**, *120*, 8907−8915.

(13) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. Fragment Molecular Orbital Method: An Approximate Computational Method for Large Molecules. *Chem. Phys. Lett.* **1999**, *313*, 701−706.

(14) Collins, M. A. Systematic Fragmentation of Large Molecules by Annihilation. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7744−7751.

(15) Dapprich, S.; Komaromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. A New ONIOM Implementation in Gaussian98. Part I. The Calculation of Energies, Gradients, Vibrational Frequencies and Electric Field Derivatives. *J. Mol. Struct.: THEOCHEM* **1999**, *461-462*, 1−21.

(16) Chesnut, D. B.; Moore, K. D. Locally dense basis sets for chemical shift calculations. *J. Comput. Chem.* **1989**, *10*, 648−659.

(17) Jensen, F. Basis Set Convergence of Nuclear Magnetic Shielding Constants Calculated by Density Functional Methods. *J. Chem. Theory Comput.* **2008**, *4*, 719−727.

(18) Jensen, F. Segmented contracted basis sets optimized for nuclear magnetic shielding. *J. Chem. Theory Comput.* **2015**, *11*, 132−138.

(19) Collins, M. A. Molecular Forces, Geometries, and Frequencies by Systematic Molecular Fragmentation Including Embedded Charges. *J. Chem. Phys.* **2014**, *141*, 094108.

(20) McConnell, H. M. Theory of Nuclear Magnetic Shielding in Molecules. I. Long-Range Dipolar Shielding of Protons. *J. Chem. Phys.* **1957**, *27*, 226−230.

(21) Stone, A. J. Distributed Multipole Analysis: Stability for Large Basis Sets. *J. Chem. Theory Comput.* **2005**, *1*, 1128−1132.

(22) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. *Gaussian 09*, Revision E.01; Gaussian, Inc.: Wallingford, CT, 2013.

(23) Reid, D. M.; Kobayashi, R.; Collins, M. A. Systematic Study of Locally Dense Basis Sets for NMR Shielding Constants. *J. Chem. Theory Comput.* **2014**, *10*, 146−152.