Pergamon

SCIENCE DIRECT®

Tetrahedron Letters 44 (2003) 8233–8236

TETRAHEDRON LETTERS

QRC: a rapid method for connecting transition structures to reactants in the computational analysis of organic reactivity

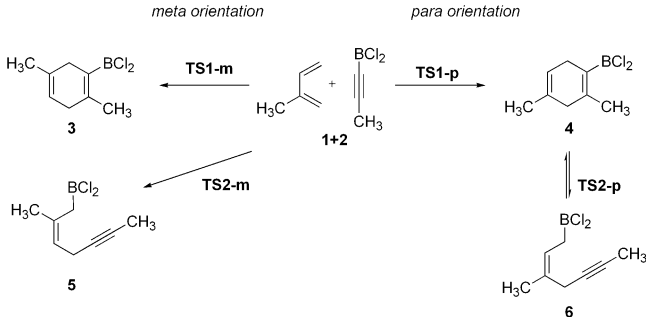
Jonathan M. Goodman* and María A. Silva

Unilever Centre for Molecular Informatics, Department of Chemistry, Lensfield Road, Cambridge CB2 1EW, UK

Received 25 July 2003; revised 2 September 2003; accepted 11 September 2003

Abstract—A new procedure, QRC calculation, is presented which provides useful information for analysing organic reactions and links transition structures to starting materials and products. QRC calculations are an alternative to computationally demanding IRC calculations, which find more information than is needed for this analysis.

In order to study organic reaction pathways, a common approach is to perform calculations on the starting materials and products and then to identify a transition structure, which may link them together. The transition structure needs to be characterised to show that it has an unique normal mode with an imaginary frequency. This is not enough to show that the transition structure links particular starting materials and products. In a recent study we identified two different but rather similar transition structures which linked different reactant pairs (Scheme 1).[1] Which reactant pair was linked by which transition state was not clear until an intrinsic reaction coordinate (IRC) calculation had been completed. IRC calculations,[2,3] which walk down the reaction coordinate to find where the minimum energy pathway leads, are often used for this purpose. High accuracy is required for the calculation of thermal rate constants,[4] and improvements in the procedure for the calculation of IRCs usually do not sacrifice precision for speed.[5] Studies of organic reactivity are trying to answer a slightly different question to the one for which IRCs were designed. For organic studies, the precise reaction coordinate is less important than the connection between the reactant and the product. Is it possible to design a procedure that follows the reaction coordinate to the starting materials and products more rapidly?



Scheme 1. *TS2-m* and *TS2-p* have similar geometries but different paths.

Here we report a procedure which does this in a fraction of the time of a full IRC calculation, but appears to identify reactants and products as effectively, at the expense of losing some detail about the precise reaction pathway. For studies of organic reactivity, the information that is lost is information that is not usually used in the first place. We call this new process *QRC* for quick reaction coordinate.

Studies of organic reactivity that identify transition structures usually report that the structures have only one imaginary frequency, but do not usually report IRC calculations. This may be because such calculations are computationally expensive, because they are complicated to set up, or because this procedure is not available. The QRC procedure alleviates these problems. It comprises two programs. The first, *FreqQRC*, takes the result of the normal mode calculation, and produces two structures, corresponding to a short step along the normal mode forwards and backwards. By default, the normalised displacements from the frequency calculation are used, but these can be adjusted by using a scaling factor after the FreqQRC command.

* Corresponding author. Tel: +44 1223 336434; fax: +44 1223 336362; e-mail: j.m.goodman@ch.cam.ac.uk

0040-4039/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tetlet.2003.09.074

This process is almost instantaneous. The second step is to minimise each of these structures, and monitor the results with the second program: *FindQRC*.

This procedure has the potential problem that the minimisation may fall into the wrong valley in the potential energy surface, and so give misleading results. This concern is probably the reason why similar approaches have not been generally used. *FindQRC* is a safeguard against this, because it allows the energy changes to be plotted against the geometry changes for the structures as the minimisation progresses, generating a profile which correlates closely with the IRC results (Fig. 1). The variations in the gradient of this line highlight the areas where the potential energy surface flattens out and then becomes steeper again. If the QRC shows difficult regions of this type, it may be desirable to do a full IRC calculation. However, we have found no examples of reactions where the QRC and IRC connect different reactants and transition structures.

The Java source code and compiled class file are both available on http://www.ch.cam.ac.uk/MMRG/software/ the program will work for both *Jaguar*[6] and *GAMESS*[7] output files, and the source code is annotated so that it could be modified to deal with output from other programs.

An IRC calculation generates an energy for each step of the pathway, and each step is the same length in mass-weighted coordinates. The length of each step is chosen so that it is likely to be smaller than the curvature of the potential energy surface. In order to compare the minimisation process with an IRC, we need to calculate some quantity that may be regarded as equivalent to the IRC step size. For this, we used the mass-weighted root-mean-square difference between structures, after aligning them using a Java translation of the quaternion fit program, *qtrfit*, of David J. Heisterberg (Ohio Supercomputer Center). The program to do this is called *FindQRC* and is available on the web site. The step size could be calculated by comparing the

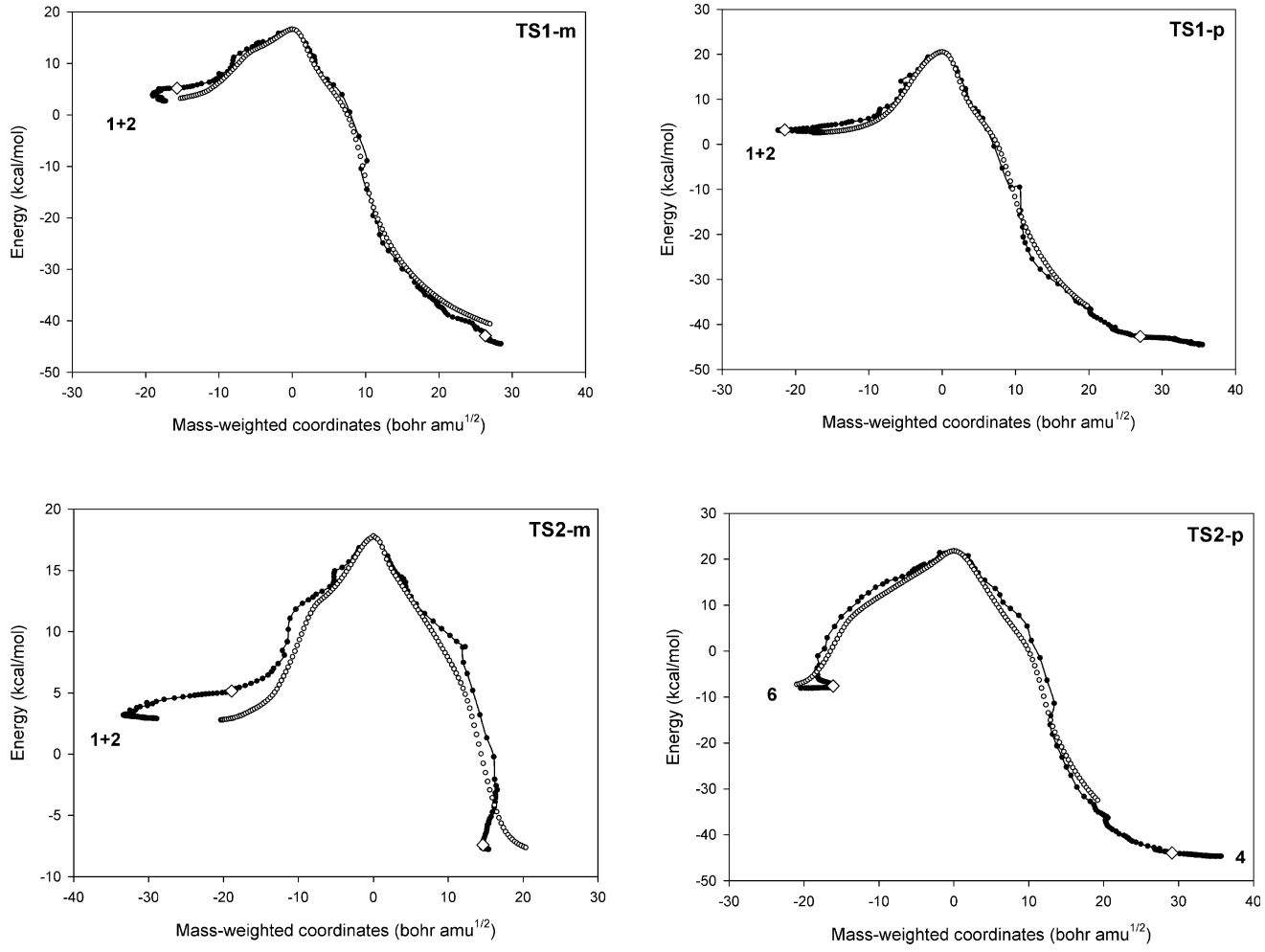


Figure 1. Plot of reaction coordinates versus energy for the transition structures in Scheme 1. (○) Refers to the IRC and (●) represents the QRC; (◇) represents where the optimisation stops when the convergence criteria is set loose (15 times the default for GAMESS). The TS2-m graph has a larger vertical scale.

Table 1. CPU times for QRC and IRC calculations[a]

TS	Steps			Time (h)		
	IRC	QRC	QRC loose[b]	IRC	QRC	QRC loose
TS1m						
Forward	90	130	70	141.4	67.2	35.8
Reverse	51	128	38	111.9	51.0	16.8
TS1p						
Forward	67	160	67	108.0	72.3	30.7
Reverse	61	114	45	121.0	54.0	22.7
TS2m						
Forward	69	99	57	186.7	46.6	27.1
Reverse	70	157	42	138.4	72.4	22.6
TS2p						
Forward	65	155	60	157.7	73.0	30.3
Reverse	70	158	73	109.0	76.0	36.3

[a] Calculations at B3LYP/6-31G* theory level.
[b] The loose convergence criteria is 15 times the default for GAMESS.

structure with the previous geometry in the minimisation process, or with the first geometry in the sequence.

Unlike an IRC calculation, minimisations will sometimes overstep and then retrace their paths. If the step size is always calculated from the previous geometry, these kinks in the path will be straightened out to produce a rather long path. We have found that calculating the distance to the original structure gives clearer results, although this occasionally means that the QRC will fold back on itself, from the measure of mass-weighted rms difference. *FindQRC* reports both measures, and omits steps in the minimisation for which the energy rises, as these are mis-steps in the minimisation. The distances are reported in Bohr amu$^{0.5}$ as is conventional for IRC calculations.

Comparisons of QRC and IRC calculations show that they produce rather similar paths. Each step of the QRC is quicker to calculate, as it corresponds to a single point calculation, and not a minimisation at a particular step size. At the beginning of the procedure, the QRC step size is bigger than the IRC step size, and so there are fewer steps to calculate. The last few steps of most minimisations are very close to each other, in terms of the mass-weighted rms-superposition of adjacent structures, so at this stage of the process, the QRC step size is smaller than the IRC step size. A loose convergence criterion can be set for the minimisation, if it is not necessary to generate a precisely minimised structure. It may well be useful, however, to use a normal convergence criterion, so that the QRC leads directly to minimised structures corresponding to the starting material or product. IRC calculations are usually set up by choosing in advance the number of steps to be taken, and the end of the IRC does not usually correspond to a minimum energy structure.

Following Schlegel,[5] we applied the procedure to the isomerisation of HCN to CNH at a low level of theory. All the calculations took less than a second, and the QRC satisfactorily reproduced the IRC. However, while the IRC calculation does not reach final geometry of HCN, QRC gave the optimised geometry for HCN. In many systems it is obvious which starting materials and products are related by a transition structure. However, there are examples where the reaction coordinate is complex or confusing. Our recent studies of the Diels–Alder reactions of alkynyl-boranes[1,8] (Scheme 1), identified two transition states, *TS1-m* and *TS2-m*, which have similar geometries and yet link different starting materials and products. This is a good test of the QRC procedure. A comparison of the IRC and QRC for these four pathways is shown in Figure 1.

In all cases, the shape of the IRC pathway is reproduced by the QRC and they connect the same species at both sides of the transition structures. A comparison of the time required for the IRC and QRC calculations is given in Table 1; it can be seen that even when the number of steps is larger in QRC, the CPU time is still considerably less than time required for the IRC calculations. The amount of CPU time for the QRC is substantially reduced by giving a looser convergence criterion.

In conclusion, the programs *FreqQRC* and *FindQRC* provide a quick and convenient method to link transition structures and reactants, and produce a plot that can show problem areas in the potential energy surface along the pathway. We are searching for examples where the QRC procedure disagrees with an IRC calculation. So far, the QRC procedure has proved effective for all the examples we have investigated.

Acknowledgements

M.A.S. thanks the European Community for a Marie Curie Fellowship.

References

1. Silva, M. A.; Pellegrinet, S. C.; Goodman, J. M. *J. Org. Chem.* **2003**, *68*, 4059–4066.
2. Fukui, K. *Acc. Chem. Res.* **1981**, *14*, 363–368.
3. Müller, K. *Angew. Chem., Int. Ed. Engl.* **1980**, *19*, 1–13.
4. Baldridge, K. K.; Gordon, M. S.; Steckler, R.; Truhlar, D. G. *J. Phys. Chem.* **1989**, *93*, 5107–5119.
5. Gonzalez, C.; Schlegel, H. B. *J. Chem. Phys.* **1989**, *90*, 2154–2161.
6. Jaguar version 4.2: Schrodinger, Inc., Portland, Oregon, 2000.
7. GAMESS version 20: Iowa State University.: Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. I.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comp. Chem.* **1993**, *14*, 1347–1363.
8. Silva, M. A.; Pellegrinet, S. C.; Goodman, J. M. *J. Org. Chem.* **2002**, *67*, 8203–8209.