

Janocchio – a Java applet for viewing 3D structures and calculating NMR couplings and NOEs

David A. Evans, Michael J. Bodkin, S. Richard Baker and Gary J. Sharman*

Eli Lilly and Company Ltd, Lilly Research Centre, Windlesham, Surrey, GU20 6PH, UK

Received 13 December 2006; Revised 26 March 2007; Accepted 2 April 2007

We present a Java applet, based on the open source Jmol program, which allows the calculation of coupling constants and NOEs from a three-dimensional structure. The program has all the viewing features of Jmol, but adds the capability to calculate both H–H and H–C 3-bond couplings constants. In the case of H–H couplings, the Altona equation is used to perform this. The program also calculates NOEs using the full relaxation matrix approach. All these calculations are driven from a simple point and click interface. The program can calculate values for multi-structure files, and can produce input files for the conformational fitting program NAMFIS. Copyright © 2007 John Wiley & Sons, Ltd.

KEYWORDS: J-coupling; NOE; Java; applet

INTRODUCTION

NMR is arguably the most powerful tool in the study of the three-dimensional structure of molecules, both in terms of distinguishing stereochemistry and in terms of determining conformation. The two most widely used NMR parameters that find use in the resolution of such structural problems are undoubtedly scalar coupling constants and nuclear Overhauser enhancements (NOEs). Karplus first discovered the relationship between 3-bond scalar or J-couplings and dihedral angles. Since then, many 'Karplus' equations have been derived, which relate couplings to dihedral angle, most notably, in the case of H–H couplings, by Altona *et al.*¹ A similar general equation, taking into account the nature of the bonded atoms, does not exist for H–C couplings, though a number of expressions relating such couplings to dihedral angle have also been proposed, such as that due to Wasylichen.^{2–5} Similarly, NOEs have found widespread use in structural studies, and many excellent reviews on the subject are available.⁶

There are many computer programs available that calculate coupling constants, such as the MestrelJ,⁷ and there are even a number of websites that perform such calculations.⁸ However, these programs do not provide a 3D structure viewing tool. There are also a number of packages for calculating NOEs, such as NOEPROM,⁹ which uses a full relaxation matrix approach. This includes a 3D viewing tool, but cannot handle J-couplings. There are of course many packages aimed at protein NMR that calculate couplings, NOEs and provide viewing capabilities, but these are not aimed at working with small molecules and can be difficult to use for such applications. Thus, despite the

fact that NOEs and couplings play such a vital role in structural studies, there is, to our knowledge, no simple, easily available computer program for small molecules that allows the display of 3D structures and the calculation of NOEs and couplings from them.

It was therefore decided to create a Java application and applet, 'Java NOe and Coupling Calculator with Handy Interactive Operation' (Janocchio), using the open source libraries of the molecular viewer Jmol¹⁰ and the Chemical Development Kit (CDK).¹¹ It aims to provide a simple and intuitive way to calculate both the NOEs and couplings. It allows the user to label atoms and choose the measurements to display with simple mouse clicks, calculates 3-bond couplings for both H–H and H–C dihedrals, and calculates NOEs using the full relaxation matrix approach. In addition, it retains the Jmol viewer interface, enabling full display options for the three-dimensional structure. It is typically run as an applet, embedded in a web page, but can also be launched as a separate desktop application.

RESULTS AND DISCUSSION

The applet interface

Janocchio features a simple point and click interface. The applet consists of four parts: the menus at the top, the molecular viewing area on the left and the coupling constant and NOE tables on the right (Fig. 1). The applet uses Jmol as the underlying viewer, so most of the capabilities and features of the Jmol package are retained. Importantly, this means there is compatibility with a number of common 3D file formats, the ability to change the molecular representation from wireframe to stick model to CPK, and full control of viewing angle and zooming via the mouse: indeed a wide gamut of the features expected in a molecular viewing package are present. The applet can be detached from the web page into a separate resizable

*Correspondence to: Gary J. Sharman, Eli Lilly and Company Ltd, Lilly Research Centre, Erl Wood Manor, Sunninghill Road, Windlesham, Surrey, GU20 6PH, UK.
E-mail: sharman_gary_gs@lilly.com

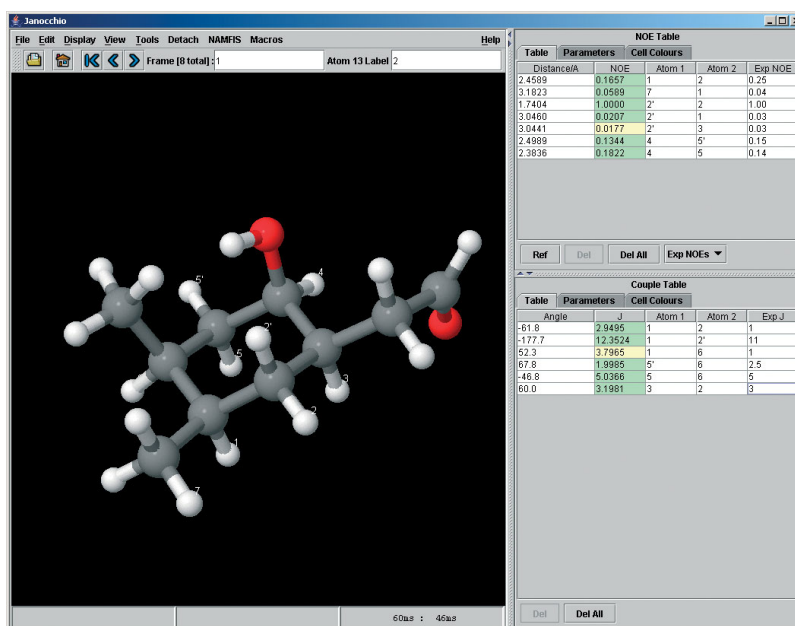


Figure 1. The main interface of Janocchio.

window. Closing this window returns the applet to the web page.

The data tables on the right show the predicted NOEs or coupling constants, along with the labels of the atoms between which the prediction is made. There is also a column for entering the experimental values as measured from NMR spectra. To facilitate a visual comparison of these values, a simple 'conditional formatting' colour coding scheme with three levels identifies values, which fit well (green), are suspect (yellow), or are plainly wrong (red). The thresholds for these levels are controllable by the user. Data from the tables can easily be copied and pasted to other applications.

A frequently occurring problem when dealing with NMR data and computer-based molecular models is that the numbering system used by the spectroscopist rarely matches that of the modelled structure. Typically, a spectroscopist uses an arbitrary numbering system, as it is rarely feasible to decide on numbering by a rule such as those of IUPAC. Although one might argue that further problems could be avoided by the spectroscopist using the numbering system of the molecular model, this is, in our experience, not the normal order in which events happen. A mapping procedure of the two numbering systems then ensues, with inevitable opportunities for errors. We have therefore implemented a feature that allows the atom labels to be changed, so as to match those used by the spectroscopist in assignment. These labels are assigned by clicking an atom to select it, and then typing the desired label in the label text field (Fig. 2). Labels appear on the structure and in the data tables. If a label is not provided, the default structure numbering is used.

A further issue with applets is that they are easily closed and their state is not persistent. We have therefore implemented a feature to save the current state of the applet, including defined measurements, labels and the presently loaded structure file, automatically to file on closing the browser or navigating to another page. The user is prompted, if they wish, to recover this session when the applet is

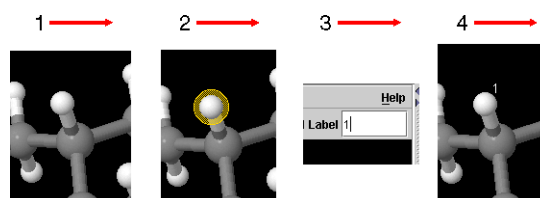


Figure 2. Labelling an atom in Janocchio allows the user to apply an existing label scheme to the structural model, avoiding the need for error prone mapping of labels by the user.

reopened. The applet state can also be saved manually to a file for later recall. The applet uses the JSON file format for this purpose.¹²

Janocchio requires the user to supply two things: NMR data (NOEs or derived distances, and couplings) and a file containing three-dimensional structures. How the user prepares this data is beyond the scope of this paper, though the quality of the output is obviously related to the quality of the input.

Coupling constants

To calculate a coupling, the user double clicks the first atom, clicks on the two intervening atoms that define the torsion, and finally double clicks on the last atom (Fig. 3). The coupling is added automatically to the table on the right side of the screen. Proton–proton couplings are calculated using the well-known Altona equation,¹ which takes into account the nature and electronegativity of nearby atoms other than the coupling partners. We have employed the recently published Java CDK libraries¹¹ to enable automatic classification of the atoms that surround a torsion, thus facilitating this calculation. Proton–carbon couplings are derived using the Wasylichen equation,² or the equations due to Tvaroska³ or Aydin.⁴ Measurements are synchronised between the molecule viewer and the tables, so every time a coupling is defined on the structure, the table is updated.

If experimental values are entered, the background of the predicted value table cell is automatically colour coded as described above, to enable easy visualisation of any error.

NOEs

Calculating an NOE is very similar to calculating a coupling constant: the user simply double clicks the two protons of interest. NOEs are calculated using the full spin relaxation matrix method.¹³ The relaxation matrix is diagonalised using a Jacobi rotation scheme, and the resultant NOE matrix calculated. Thus, all possible NOEs are calculated: the act of selecting atom pairs in Janocchio determines which are displayed in the NOE table. The calculation is performed 'on the fly' as a structure is loaded into Janocchio. The calculation is extremely fast, with no perceivable delay for typical sized molecules of *ca* 30 hydrogen atoms. The predicted NOEs can be normalised with respect to a reference intensity, or with respect to the diagonal. Parameters for the relaxation matrix calculation, such as field strength, mixing time and correlation time, are adjusted from the 'Parameter' tab (Fig. 4). A switch also allows the calculation for a 2D NOESY or 2D ROESY experiment. Typically, we do not measure correlation times but use an estimated value of 40–80 ps. A simple procedure to refine this, if required, is to adjust the correlation time until the size of the calculated NOE between two protons at a fixed distance, referenced to the diagonal, matches that of the experimental spectrum. Janocchio assumes the same correlation time for the whole molecule.

As with the coupling constants, measurements are synchronised between the molecule viewer and the tables. Likewise, when the experimental values are entered, the background of the predicted value table cell is automatically

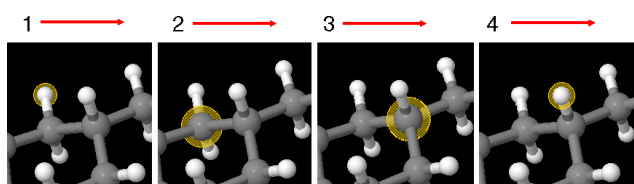


Figure 3. Defining a coupling constant is simply achieved by clicking on the atoms of the torsion. The atoms are conveniently highlighted. NOEs are defined in a similar way.

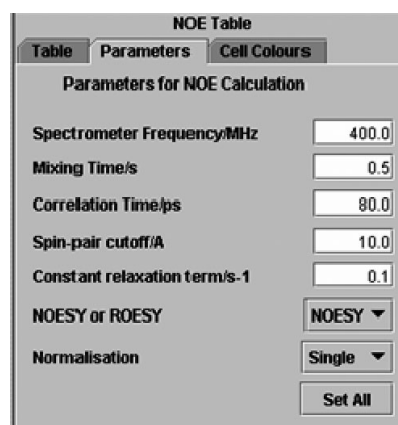


Figure 4. The parameter tab allows adjustment of the parameters for the NOE relaxation matrix calculation.

colour coded to enable easy visualisation of any error. The NOE table can be used in two ways. Experimental NOEs can be entered to compare with those calculated as described above; or, if the user prefers, experimental distances, derived for example from 1D NOE experiments at one or more mixing times, can be entered. Here, the comparison is with the distance from the model; that is, the NOE calculation is not used. Methyl groups are automatically detected and handled in the simulation using Tropp averaging; atoms given the same label are treated as equivalent and a $1/r^6$ averaging used; otherwise the assumption is that all the protons are unique.

Multi-conformer files

One of the aims of Janocchio was to enable easy comparison of different possibilities for conformations or stereochemistry of a structure. The ability to read files containing multiple conformers was therefore included. When a file containing more than one conformer is loaded, the structures can be stepped through with the frame arrow keys on the menu bar. As each structure is displayed, labels and measurements first applied to one conformer are automatically applied to the others (Fig. 5). NOEs and couplings are recalculated for the new structure and the tables updated. If experimental values have been entered, the conditional formatting is also updated. Thus, it is easy to step through a number

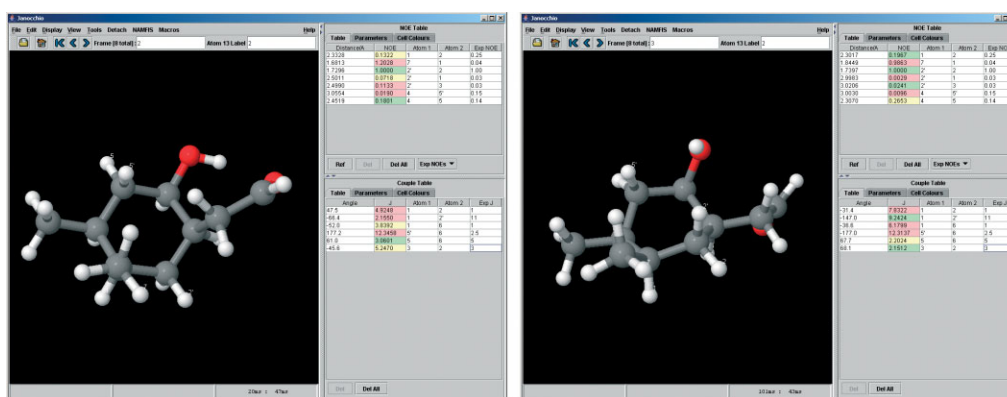


Figure 5. Stepping through the structures of a multi-conformer file automatically updates the NOEs and couplings to the newly displayed structure.

of structural possibilities and compare visually, which one best fits the experimental NMR data, by looking for the structure with the minimum number of red or yellow cells in the coupling and NOE tables.

NAMFIS input and output files

The NMR analysis of molecular flexibility in solution (NAMFIS)¹⁴ program estimates the relative population of different conformers in solution, on the basis of experimental distances derived from NOEs and coupling constants. Apart from visualising individual conformers separately, Janocchio can also generate NAMFIS input files, once a multi-conformer file has been read in and experimental inter-proton distances and coupling constants have been entered. In addition, it can read in a NAMFIS output file and display the populations derived by NAMFIS against each structure, for simple visualisation of the population analysis result.

An example of the application of Janocchio to a structural problem

We have found Janocchio to be a valuable tool in our day-to-day work with stereochemical and conformational problems. A typical example, with an interesting chemical twist, is described here. A simple deprotection reaction had been attempted on a pair of diastereomeric compounds, (1) and (2) (Fig. 6). Reaction product from one of the compounds gave an NMR spectrum consistent with the expected product, but the product from the second compound gave an NMR spectrum, which although on first inspection appeared consistent with the desired product, did not stand up on closer scrutiny. Although the correct number of protons with approximately the correct chemical shifts was present, the COSY did not show the expected coupling partners. In particular, proton 5 appeared to be coupled to proton 3, and proton 6 showed no coupling to proton 5 (Fig. 7). Furthermore, the carbon shift of carbon 5, as revealed by the HSQC, was not consistent with the expected structure. After further analysis of all of the NMR data, the weight of evidence suggested that the structure was instead (3) (Fig. 8). This structure has two possible diastereomers (3a and 3b), discounting a highly strained *trans* ring junction. We wanted further evidence that this structure was correct, and wanted to distinguish

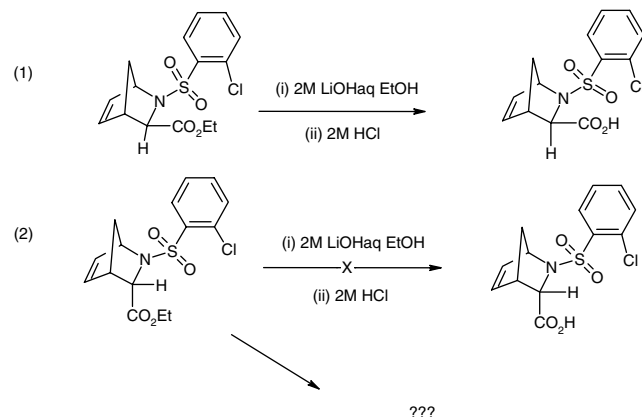


Figure 6. Structures of the two diastereomers whose basic ester hydrolysis followed by acid work-up lead to an unexpected product.

which isomer was present. Demonstrating that the coupling constants and NOEs were consistent with one of these structures would seem an ideal way to do this, and Janocchio provides us with a simple way of visualising this.

The NOEs from the NOESY spectrum were therefore quantified by integration, and proton–proton and proton–carbon coupling constants were measured. Both the possible diastereomers were modelled. The protocol followed was to run a conformational search, followed by clustering of the resulting structures. One representative conformer from each diastereomer was then viewed in Janocchio, and the NOEs and couplings measured from the NMR spectra were entered for comparison.

The results for diastereomer 3a show generally a good fit across the NOEs and couplings, suggesting the core of the suggested structure is correct (Fig. 9). However, there are a few notable exceptions, such as the coupling from proton 2 to proton 3, which is predicted at 3.2 Hz, but measured as 8.8 Hz, a very large error. Similarly, the coupling of proton 2 to carbon 7 is predicted at 3.1 Hz, but measured as 0.5 Hz. The results from the NOE table show similar patterns, with a generally good fit but with a number of significant discrepancies. In particular, the NOEs between atoms 2 and 3, and 2 and 7 appear rather larger than expected, and the NOE between atoms 2 and 4 appears much smaller than expected. As all of these errors involve atom 2, it seems highly likely that 3a is the wrong diastereomer. If the results for 3b are now considered (Fig. 10), it is clear that there is an excellent fit and the structure of the compound is confirmed.

A mechanism for this rearrangement, involving attack of the double bond by the carboxylic acid group, can easily be proposed to rationalise the formation of this product (Fig. 11). Indeed, a search for similar structures in the literature revealed that this rearrangement had actually been exploited synthetically in the past.¹⁵ Furthermore, the stereochemistry

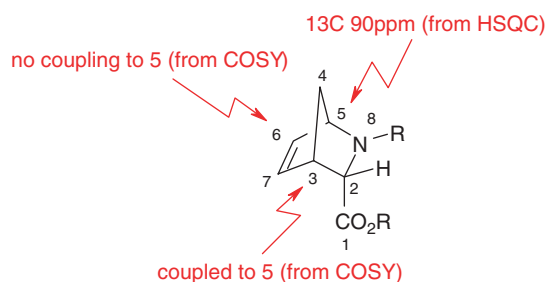


Figure 7. Inconsistencies in the NMR spectra, which suggested that the deprotection of (2) had not proceeded as expected.

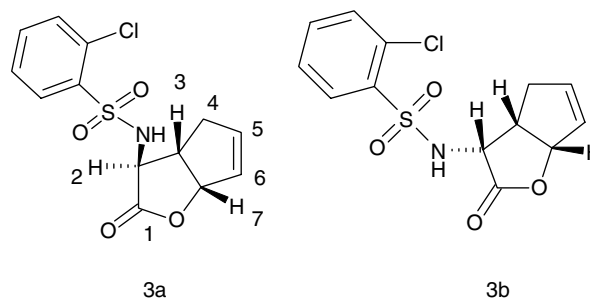


Figure 8. Possible structures of the unexpected product (3).

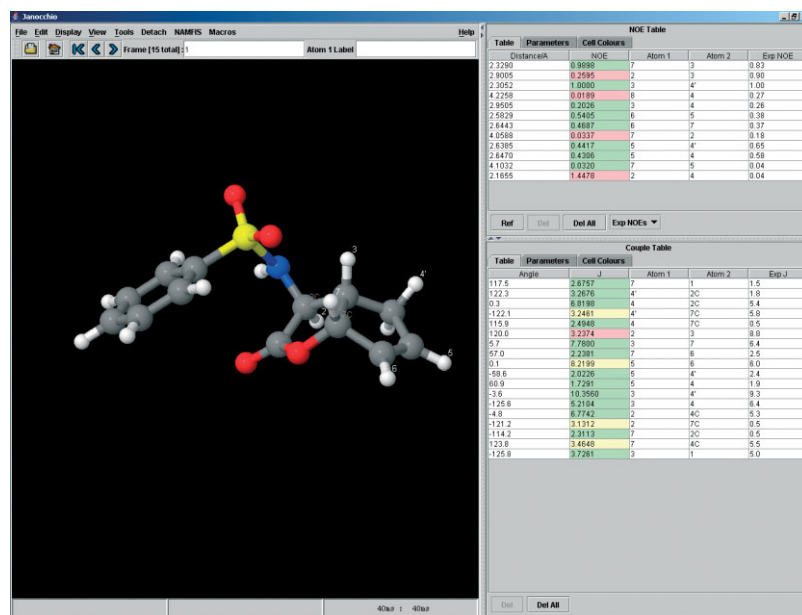


Figure 9. View of a representative conformer of structure (**3a**) in Janocchio showing match of the experimental and calculated coupling constants and the NOEs.

derived above is explained by this proposed mechanism, as it proceeds with retention of the stereochemistry present in the original molecule. Likewise, the fact that the *exo* isomer does not undergo this rearrangement is readily explained, as this isomer does not have the required geometric configuration to allow attack of the double bond by the carboxylic acid group.

Of course, one does not need Janocchio to come to the above conclusions. A spectroscopist could consider the various pieces of information and arrive at this conclusion without the assistance of the applet. However, it does provide a convenient and simple way of visualising all the data concurrently, and readily indicates apparent errors or areas of a structure that do not fit the data. Of course, using Janocchio does not take away the requirement to understand limitations and errors involved in such comparisons: over-interpretation could easily lead to erroneous conclusions.

EXPERIMENTAL

Janocchio was developed using version 5 of the Netbeans IDE.¹⁶ The applet is written entirely in Java, and as such requires that a Java run time environment (JRE) be installed. We have used the applet mainly with the Sun Microsystems JRE (version 1.4.2) for Microsoft Windows, but in principle, it will run on any platform for which a JRE of version 1.4 or above is available.

Janocchio is distributed as a 'jar' file, a Java archive that contains the class files necessary to run the program. The jar archives for Jmol,¹⁰ CDK,¹¹ the Netbeans custom layout¹⁶ and JSON file format,¹² are also required. Janocchio is not normally run as a separate application, but is accessed as an applet via the Java plug-in of an internet browser. All the jar files must be signed to enable the applet to read and write files on the disk of the computer, where the browser is running. To run the applet, the files are called from a web page with the following HTML code:

```
(APPLET id='nmrapp' width=820 height=610 align='baseline' code='org.openscience.jmol.app.NmrApplet.class' archive='Jmol_Signed.jar, Janocchio_Signed.jar, cdk-20050826_Signed.jar, swing-layout-1.0_Signed.jar, json-Signed.jar')
```

```
(/APPLET)
```

If the user has the jar files locally, however, it can be run as an application by launching the Janocchio.jar file.

All NMR experiments were performed on a Bruker DRX600 spectrometer equipped with a TXI cryoprobe. The temperature for the experiments was 298 K, and the sample was dissolved in CDCl₃ at an approximate concentration of 5 mg ml⁻¹. Proton, COSY, HSQC, HMBC, NOESY, HETLOC and phase sensitive HMBC experiments were acquired. For the proton experiment, a spectral width of 14 ppm was used and 32 K complex data points were collected. The pulse sequences for COSY, HSQC and HMBC were standard Bruker library sequences, acquired with 2 K data points over an 8 ppm spectral width. The NOESY employed zero quantum purging and random order acquisition, but was acquired with the same number of data points as described above. The mixing time used was 0.5 s, and the zero mixing time extrapolation method used to calculate the final distance values. The HETLOC¹⁷ sequence used was that of Uhrin *et al.*¹⁸ The experiment was run with a spectral width of 8 ppm. 8 K points were acquired for each of the 128 increments. The phase sensitive HMBC sequence was as described by Keeler and Edden.¹⁹ For this experiment, 8 K points were acquired in f2, with 128 increments. For both the HETLOC and phase sensitive HMBC, the data was zero-filled in f2 to 32 K prior to Fourier transformation: a strip transform was used on the regions of interest to restrict the size of the processed data set.

The structures of the two diastereomers were modelled using MacroModel v 9.1.²⁰ Conformational analysis was

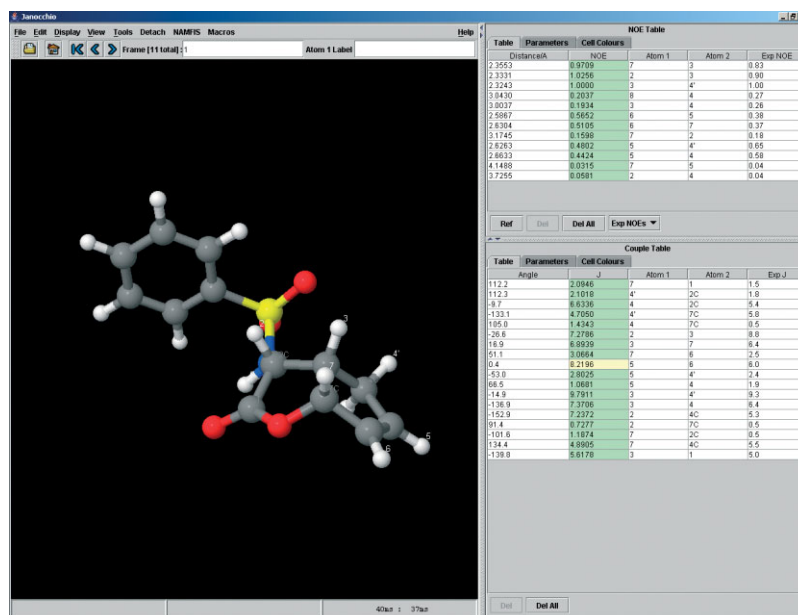


Figure 10. View a representative conformer of structure (**3b**) in Janocchio showing match of the experimental and calculated coupling constants and the NOEs.

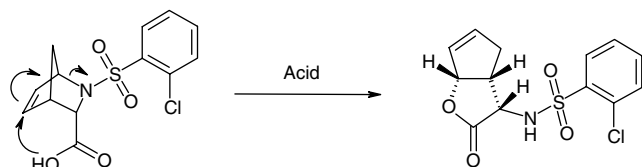


Figure 11. Possible mechanism for the formation of the rearranged product.

performed using the MMFF force field²¹ and the Monte Carlo Multiple Minimum (MCM) search protocol, minimising each found structure to a root-mean-square gradient of $<0.05 \text{ kJ mol}^{-1} \text{ \AA}^{-1}$. Clustering was performed using XCluster.²⁰ The heavy atom coordinates were used as input, one cluster of conformers was judged to be significant for each diastereomer, and the lowest energy member of these clusters was used as input for Janocchio.

CONCLUSIONS

We have built a Java applet that can be executed from within a web page, which allows prediction of NMR coupling constants and NOEs, and easy visualisation of three-dimensional structures. We believe that it provides an excellent way to bring together structural information for small organic molecules. The compiled applet, application, and source code is available from the authors or via 'sourceforge', along with a tutorial and installation guide (<http://janocchio.sourceforge.net/>).

Acknowledgements

The authors would like to thank Dr Manuel Martin-Pastor, Universidad de Santiago de Compostela, whose source code for NOEPROM was used to assist in developing the Java relaxation

matrix calculations, and Prof. Robert Hanson, St Olaf College, for adding important features to Jmol.

REFERENCES

- Haasnoot CAG, De Leuw FAAM, Altona C. *Tetrahedron* 1980; **36**: 2783.
- Wasylishen R, Schaefer T. *Can. J. Chem.* 1973; **51**: 961.
- Tvaroska I, Taravel FR. *Adv. Carbohydr. Chem. Biochem.* 1995; **51**: 15.
- Aydin R, Guenther H. *Magn. Reson. Chem.* 1990; **28**: 448.
- Conteras RH, Peralta JE. *Prog. Nucl. Magn. Reson. Spectrosc.* 2000; **37**: 321.
- Neuhaus D, Williamson MP. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*. Wiley-VCH: New York, 2000.
- <http://www.mestrec.com/> [accessed 2006].
- http://www.spectroscopynow.com/FCKeditor/UserFiles/File/specNOW/HTML%20files/General_Karplus_Calculator.htm [accessed 2006].
- Martin-Pastor M. *Estructura y dinamica de oligosacaridos mediante RMN y cálculos de mecánica y dinámica molecular (Structure and dynamic of oligosaccharides through NMR and molecular mechanics and dynamics calculations)*. Universidad Autonoma: Madrid, Spain, 1997.
- Jmol, <http://www.jmol.org/> [accessed 2006].
- Steinbeck C, Yongquan H, Kuhn S, Horlacher O, Luttmann E, Willighagen E. *J. Chem. Inf. Comput. Sci.* 2006; **43**: 493.
- JSON, <http://www.json.org/> [accessed 2006].
- Forster MJ. *J. Comput. Chem.* 1990; **12**: 292.
- Cicero DO, Barbato G, Bazzo R. *J. Am. Chem. Soc.* 1995; **117**: 1027.
- Ward SE, Holmes AB, McCague R. *Chem. Commun.* 1997; **21**: 2085.
- Netbeans 5.0, <http://www.netbeans.org/> [accessed 2006].
- Kurz M, Schmieder P, Kessler H. *Angew. Chem. Int. Ed.* 1991; **30**: 1329.
- Uhrin D, Batta G, Hruby VJ, Barlow PN, Kover KE. *J. Magn. Reson.* 1998; **130**: 155.
- Edden RAE, Keeler J. *J. Magn. Reson.* 2004; **166**: 53.
- MacroModel, Schrodinger. LLC: New York, 2005.
- Halgren TA. *J. Comput. Chem.* 1996; **17**: 490.