CAL3JHH: a Java program to calculate the vicinal coupling constants ($^3J_{\mathrm{H,H}}$) of organic molecules

Alonso Aguirre-Valderrama · José A. Dobado

Abstract Here, we present a free web-accessible application, developed in the JAVA programming language for the calculation of vicinal coupling constant ($^3J_{\mathrm{H,H}}$) of organic molecules with the H–C$_{\mathrm{sp3}}$–C$_{\mathrm{sp3}}$–H fragment. This JAVA applet is oriented to assist chemists in structural and conformational analyses, allowing the user to calculate the averaged $^3J_{\mathrm{H,H}}$ values among conformers, according to its Boltzmann populations. Thus, the CAL3JHH program uses the Haasnoot–Leeuw–Altona equation, and, by reading the molecule geometry from a protein data bank (PDB) file format or from multiple pdb files, automatically detects all the coupled hydrogens, evaluating the data needed for this equation. Moreover, a "Graphical viewer" menu allows the display of the results on the 3D molecule structure, as well as the plotting of the Newman projection for the couplings.

Keywords 3JHH · Coupling constant · NMR · Haasnoot equation · JAVA

Introduction

The nuclear magnetic resonance (NMR) technique[1] is nowadays a powerful tool to elucidate the structure of organic molecules. Moreover, the vicinal proton–proton coupling constants[2] ($^3J_{\mathrm{H,H}}$) contribute strongly [1] to this

effort (conformational analysis of macromolecules), since its dependency was established with the dihedral angle between the coupled hydrogens by means of the Karplus equation [2]. For example, the utility of this approach has been proven in the conformational analysis of a six-membered puckered-ring with the substituents in the axial/equatorial arrangements [3].

In addition, to include the influence of additional molecular parameters (e.g., the electronegativity and the relative position of the substituents bonded to the H–C–C–H fragment) different Karplus-type equations were developed[3] [4]. Among those equations, the most popular one is the Haasnoot, Leeuw and Altona (HLA) [4]:

$$
\begin{aligned}
^3J_{\mathrm{H,H}} = &\, P_1 \cos^2 \phi + P_2 \cos \phi + P_3 \\
&+ \sum \Delta\chi_i [P_4 + P_5 \cos^2(\xi_i \phi + P_6 |\Delta\chi_i|)]
\end{aligned}
\tag{1}
$$

where the sum extends to all the α-substituents attached to the H–C–C–H fragment. Moreover, χ is the group electronegativity for the α-substituents corrected with the electronegativity of the β-substituents, and defined as:

$$
\Delta\chi^{\mathrm{group}} = \Delta\chi^{\alpha-\mathrm{substituent}} - P_7 \sum \delta\chi^{\beta-\mathrm{substituent}}
\tag{2}
$$

where $\Delta\chi_{\mathrm{atom}} = \chi_{\mathrm{atom}} - \chi_{\mathrm{H}}$, with χ being the Huggins electronegativity [6]. The ξ_i parameter accounts for the orientation of the i-substituent (with $+1$ or -1 values). P_1 to P_7 are empirical parameters determined from 315

A. Aguirre-Valderrama · J. A. Dobado (✉)
Grupo de Modelización y Diseño Molecular, Departamento de Química Orgánica. c/ Severo Ochoa s/n, Universidad de Granada, 18071 Granada, Spain
e-mail: dobado@ugr.es

---

[1] Richard R. Ernst in 1991 and Kurt Wüthrich in 2002 were awarded with the Nobel price by their contributions to the development of this technique. See http://www.nobel.se for more information.

[2] The $^3J_{\mathrm{H,H}}$ is defined as the splitting between the ^{1}H chemical shifts peaks, originated by the additional contributions of the magnetic fields of vicinal hydrogens to the magnetic field of each hydrogen.

[3] See [5] and references therein.

reference coupling constants, and the ϕ symbol corresponds to the H–C–C–H dihedral angle.

This equation is frequently used in conformational-analysis studies [7], enabling the validation of the computed molecular geometries by comparison with the experimental data. Some specific computer applications [8–10] and molecular-modeling programs[4] utilize this equation. However, there is another significant aspect, in the conformational-analysis field, which has still not been considered in such programs: the experimental NMR data are collected normally in a liquid solution of the sample, and commonly within an equilibrium of different conformations, corresponding the experimentally measured couplings to the averaged values for all the conformations that contribute to the equilibrium.[5]

In this work, we present a JAVA application which allows calculations of the $^3J_{H,H}$ constants for one or multiple molecules (with H–C_{sp3}–C_{sp3}–H fragments). Moreover, the CAL3JHH program computes the averaged $^3J_{H,H}$ values for the different conformers, when the energies are supplied. Thus, the program uses the Boltzmann distribution [12] (i.e., the most probable distribution of conformers according to the energies) from the Statistical Thermodynamics to determine the conformers population:

$$\frac{N_i}{N} = \frac{e^{-E_i/kT}}{\sum_i e^{-E_i/kT}} \qquad (3)$$

where E_i is the energy (Kcal/mol) of each conformer, k is the Boltzmann constant, T is the temperature (Kelvin) and N_i/N is the probability of finding the conformer with a fixed energy value E_i, (i.e., the ratio of each conformer).[6]

Furthermore, like other software developed by our group [13], CAL3JHH is accessible free of charge via internet from our web page [14], avoiding the difficult distribution process to maintain a multiplatform software with different installations. With this aim, we chose the JAVA programming language [15] and implemented the program as an applet. The user needs only a web browser to run it without worrying about future upgrades or installations.

---

[4] PCMODEL. Available on Serena Software, PO Box, 3076, Bloomington, IN 47402-3076, U.S.A.

[5] An illustrative case are the carbohydrates compounds which in liquid solution give mixture of open-chains and puckered-rings with different structures (e.g., the pyranose ring is a mixture of 38 different canonical ring forms, plus the OH hydroxyl groups which yield different rotamers). See for example: [11].

[6] The program computes the couplings of each conformer and those values are affected by the ratio in a proportion multiplied by one. Being the final averaged value for each constant, the sum of all the particular values multiplied by its ratio.

Computational methods

The Java(TM) SDK 6 Standard Edition[7], v1.6.0_06, from Sun Microsystems [16] was used to develop the program. Java is an Object-Oriented Programming (OOP) language of which the main characteristics are the platform independency and ability to run in any web browser as an applet. An applet is a program which is downloaded from a HTML page and executed in the browser, provided that the browser supports the Java Virtual Machine (JVM). The CAL3JHH has been tested for MS-Explorer v6.0 and Netscape v7.0.

Performing the different reading operations from PDB files [17], the program needs reading permission on the local machine, due to the strict security of JAVA. The JAR archive containing the applet is therefore signed and the user is requested this permission before executing the program.

The CAL3JHH program implements the aforementioned HLA equation, and the different sets of P_1 to P_7 parameters were taken from the original work [5]. The HLA equation has r.m.s. error of 0.48 Hz, for the $^3J_{H,H}$ constants.
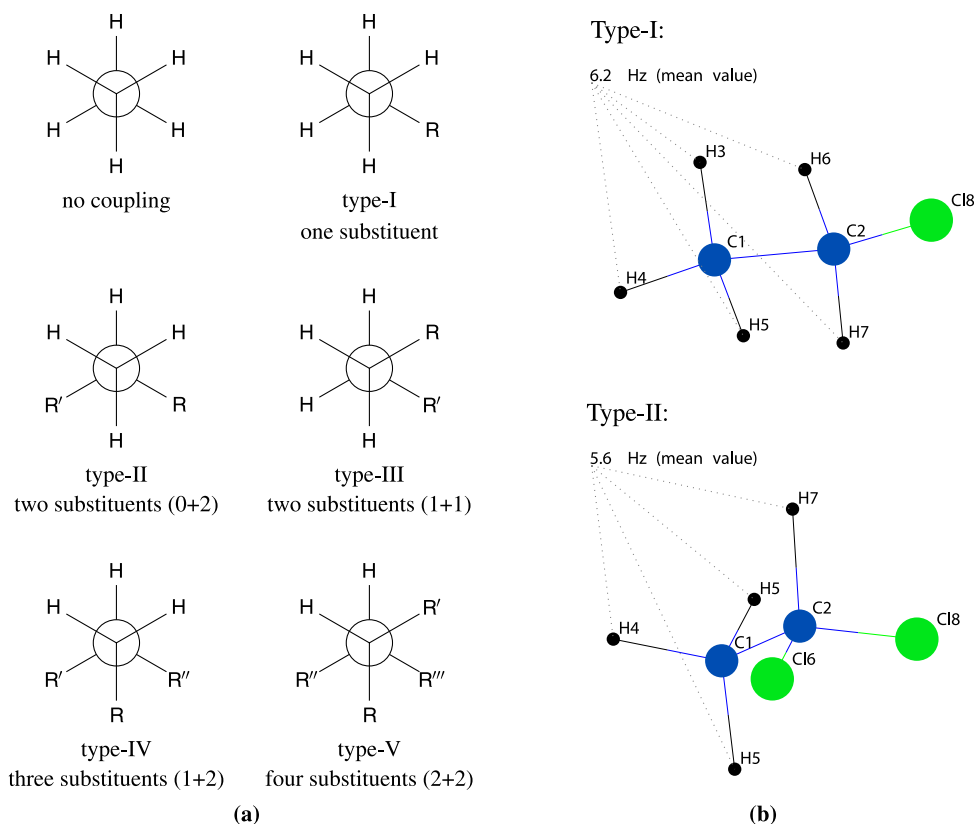
The program automatically detects all the coupled H–C_{sp3}–C_{sp3}–H fragments from the PDB file. The required information (atom type, coordinates and connectivity) are read from the protein data bank (PDB) file and, from those, all the necessary information (dihedral angles, electronegativity of the substituents and their relative positions) is processed to achieve the calculation of the $^3J_{H,H}$ values. If connectivity is not present, the program guesses the connectivity from the atomic distances, compared with the covalent radii.

The strychnine molecule and the 1,1,2-trichloropropane conformers used for the "Compute multiple structures" menu example, were optimized at the UFF and HF/6-311G*//HF/6-311G* theoretical levels, respectively, with the GAUSSIAN program [18], and later converted in the PDB format with the Newzmat utility to be used as input to the CAL3JHH.

On the other hand, when a C_{sp3}–C_{sp3} bond presents several H–C–C–H couplings (e.g., methylene or methyl groups), the program also generates the mean $^3J_{H,H}$ values. Due to possible free rotations around the single C–C bond, in some cases, the NMR spectrometer detects the hydrogen in an averaged environment collecting only the mean value. Figure 1a depicts the different types of proton environments dependent on the substituents: for the type-I and -II the situation is the same in both case, and the couplings constants in CH_3 moieties are averaged due to free rotation. For type-III and -IV, for example in the particular case of 1,2-dichloroethane (for type-III) or

---

[7] Downloaded freeware from the JAVA web page, see reference [15].

Fig. 1 Different environments for the generic vicinal C_{sp3}–C_{sp3} couplings. Only the types-I and -II yields mean values. (**a**) Newman projection plot for the types I–V. (**b**) Graphical output with the mean $^3J_{H,H}$ values generated by CAL3JHH, for the type-I (chloroethane) and type-II (1,1-dichloroethane)

no coupling

type-I
one substituent

type-II
two substituents (0+2)

type-III
two substituents (1+1)

type-IV
three substituents (1+2)

type-V
four substituents (2+2)

(a)

Type-I:

6.2 Hz (mean value)

Type-II:

5.6 Hz (mean value)

(b)

1,1,2-trichloroethane (for type-IV) for which R = R' and R = R' = R", respectively, only one signal is experimentally obtained and the CAL3JHH gives their main values in these or similar cases. Finally, in type-V, only one constant is present, and obviously there is nothing to average.

Description of the CAL3JHH Java program

CAL3JHH uses self-explanatory menus and elements for the sake of simplicity, and it is organized in three modules arranged in a tabbed panel: "Compute single structure" menu, "Compute multiple structures" menu, and one additional "Graphical viewer" menu to display graphically the results.

"Compute single structure" menu

Figure 2 shows the program screen after a "Compute single structure" run for the strychnine molecule. With the "Open file" button a PDB file can be selected from the user local machine, and the program automatically will compute all the couplings present in the molecule. Then, the results are displayed on the text area. At the end of the output are listed the couplings computed together with their corresponding dihedral angles and couplings identifications (for

example, the following character chain is shown in the 4th row:

$$[-\text{C7}(\text{C3}, \text{C8}, \text{C12})][\text{H28}]; \text{H27}-\text{C13}-\text{C22}-\text{H37};$$
$$[\text{H38}][-\text{N21}(\text{C12}, \text{C32})]$$

where 'H27−C13−C22−H37' indicates the coupled protons of the C–C bond; each chain enclosed in brackets correspond to each α-substituent with its β-substituents in parenthesis. The minus or plus signs indicate the relative position of the α-substituent.[8]

The program supports an additional way to calculate atomic connectivity, comparing the interatomic distances with the respective covalent radii (see the "from covalent radii" menu at the bottom-right of Fig. 2), because some PDB files can be generated by other programs without the atom connectivity ('CONNECT' records[9]). In this particular case, a warning message, in the output results together with the calculated new connectivity in a PDB format, notifies the user to check it carefully for consistency.

---

[8] The sign is not listed when the substituent is hydrogen since it does not participate, because the HLA equation accounts for the relative electronegativity ($\Delta\chi_{atom} = \chi_{atom} - \chi_H$) of the substituents.

[9] The 'CONNECT' records define the bonds not included in the standard residue connectivity table.

Fig. 2 View of the "Compute single structure" menu of CAL3JHH, for the strychnine molecule
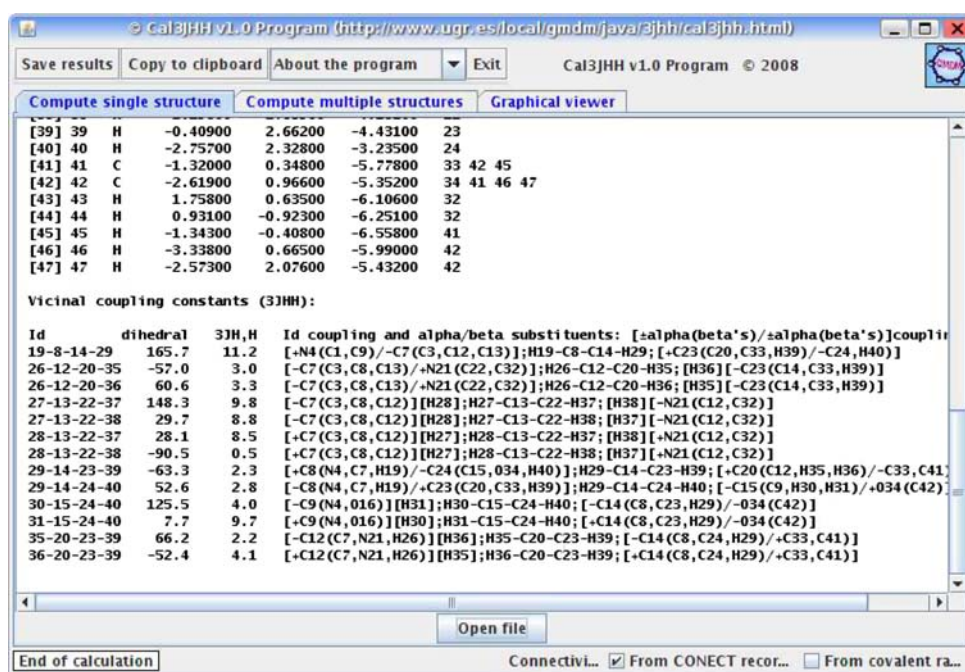
Furthermore, this program presents an additional menu (Fig. 2 up-left) that allows the user to "Save results" to a text file, "Copy results to clipboard", to find out "About" the program, and "Exit". Finally, on the bottom-left corner a label provides information on the final status of the program with text messages.

"Graphical" viewer menu

The Fig. 3 depicts the CAL3JHH "Graphical viewer" menu with a frame displaying the previous calculations of strychnine molecule (see Fig. 2). The module includes viewers for the molecule in 3D[10] and the corresponding plot in the Newman projection (where all the torsion angles involved in the constants are displayed), for the selected couplings. We developed a very simple viewer, in order to gain speed in the molecule rotation with slow PCs, resulting in a small JAR file that decreases the Internet download process. By clicking on the "Couplings" list, the user can select the different couplings from the molecule, while the numerical value is listed in the 3D panel. An additional menu (bottom-right) allows the user to show/hide the atom labels.

"Compute multiple structures" menu

Finally, as an example, Fig. 4 shows a "Compute multiple structures" run for the three gauche conformers of the 1,1,2-trichloropropane. This menu includes a table (left) of

four columns, with the list of molecules (conformers) to compute together with their energies[11], where the different molecules can be selected/unselected or Energy values edited. Molecules can be deleted/appended from this table with the bottom menus "Add molecules to table", "Clear table" and "Invert table selection".

Once the molecules to be computed are selected, the "Compute" button runs the calculations. If all the energy values are provided and all the molecules are the corresponding conformers of the same molecule, the program will calculate the $^3J_{H,H}$ averaged values, after prompting for the temperature. On the contrary, the program calculates only the $^3J_{H,H}$ values for each molecule from the list.[12]

When the averaged values are computed a final table is listed in the output with the relative energies, Boltzmann population (%) and the couplings for each conformer (see the Fig. 4 bottom).

After computing the couplings, the "Graphical viewer" displays (see Fig. 5) the corresponding result for the 1,1,2-trichloropropane. As mentioned above, the molecule can be rotated with its corresponding couplings, but now the "Molecules" list includes the different conformers that can be chosen by mouse clicking on the list.

---

[11] The energies are only needed to compute the averaged values among conformers. The input of the energy values for the different conformers can be set manually in this table—the E(kcal/mol) values in the corresponding column can be edited, or by default the energy values are read from each PDB file when a 'REMARK' record is found beginning with the text 'ENERGY(KCAL/MOL)=value', (where value means the energy data).

[12] The CAL3JHH has no restrictions on the number of molecules (or the file size) that can handle, except for the local hardware limitations.

---

[10] Rotation and zoom operations are available trough pressing the x, y, z and t letter in keyboard while dragging the mouse.

Fig. 3 Graphical viewer menu of CAL3JHH displaying the results for the strychnine molecule

Fig. 4 View of a "Compute multiple structures" menu for the 1,1,2-trichloropropane conformers

The importance of accounting the conformers population is evident after comparing the experimental and the calculated couplings. Figure 6 plots the Newman projections together with the experimental data [19] for the three aforementioned calculated conformers. As can be seen from the $J_{H5,H10}$ coupling, the calculated value agrees with the experimental one, when the three conformers are considered.

Fig. 5 "Graphical viewer" with the results for the 1,1,2-trichloropropane
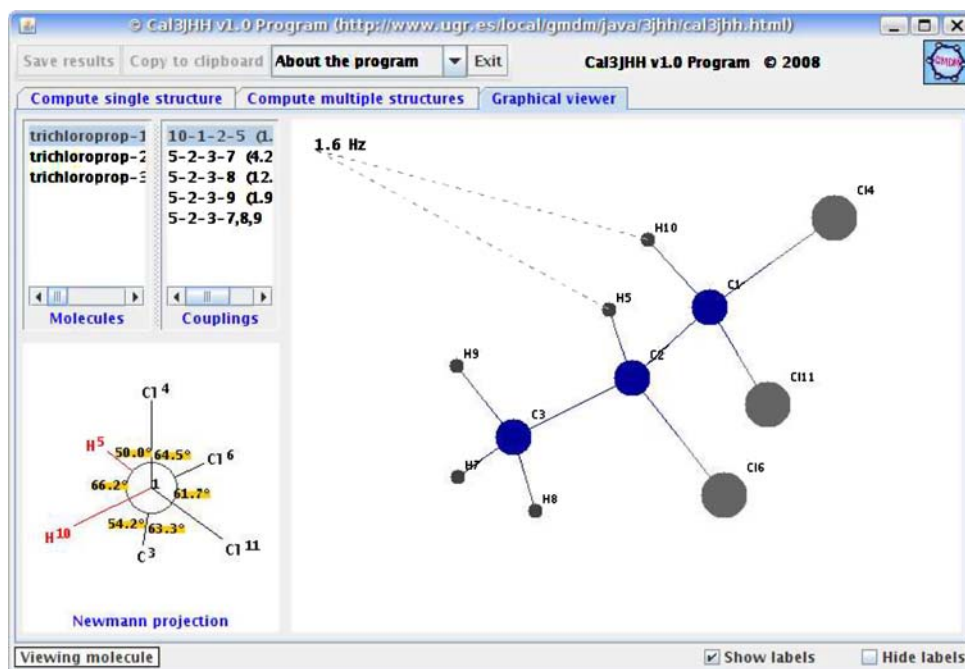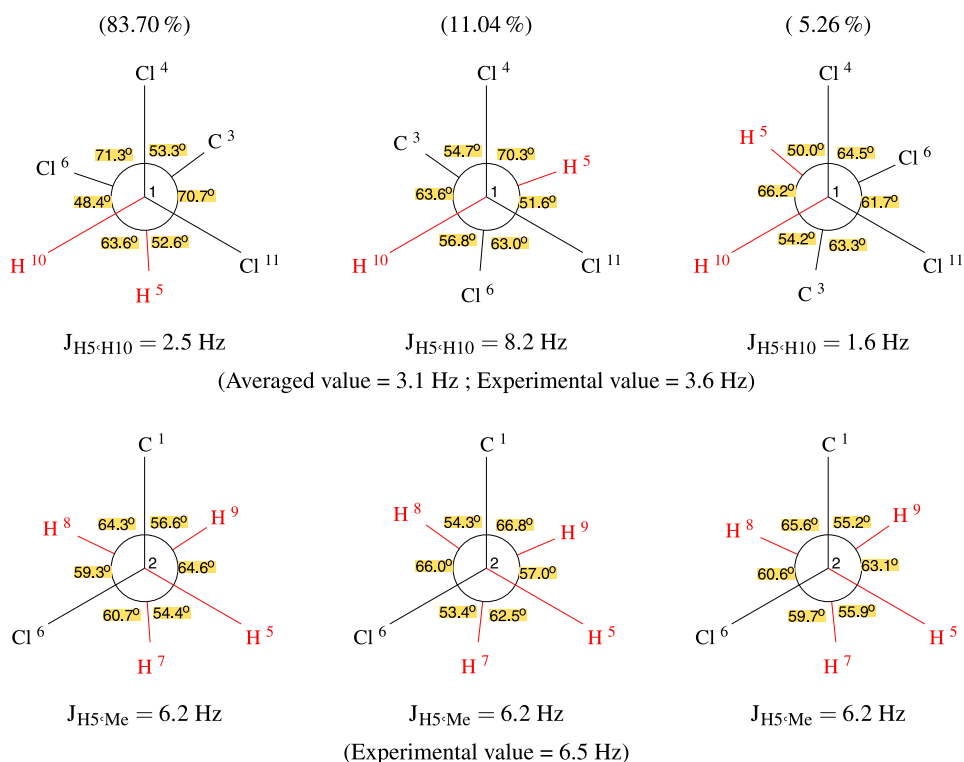

Fig. 6 Newman projection plot for the different 1,1,2-trichloropropane conformers together with their energies (experimental values from ref. [19])

(83.70 %)   (11.04 %)   (5.26 %)

$J_{H5\cdot H10} = 2.5$ Hz $J_{H5\cdot H10} = 8.2$ Hz $J_{H5\cdot H10} = 1.6$ Hz

(Averaged value = 3.1 Hz ; Experimental value = 3.6 Hz)

$J_{H5\cdot Me} = 6.2$ Hz $J_{H5\cdot Me} = 6.2$ Hz $J_{H5\cdot Me} = 6.2$ Hz

(Experimental value = 6.5 Hz)

Additional examples

In order to highlight the usefulness of the CAL3JHH program, in this section is included two additional illustrative examples.

In the first example (see Fig. 7) we show the screenshot of the "Graphical Viewer" with the results of the $^3J_{H,H}$

calculations of the α-fucosidase protein (unit A) [20].[13] The protein (unit A), with all the non-polar hydrogens added,

_____

[13] Although, the protein backbone has no H-Csp3-Csp3-H torsion angles, and the calculation of vicinal coupling constants are applicable only to the substituents, we include this example to illustrate the power of calculation of CAL3JHH program to automatically recognize such coupling constants for large PDB files.

Fig. 7 "Graphical viewer" with the results for the unit A of the α-fucosidase
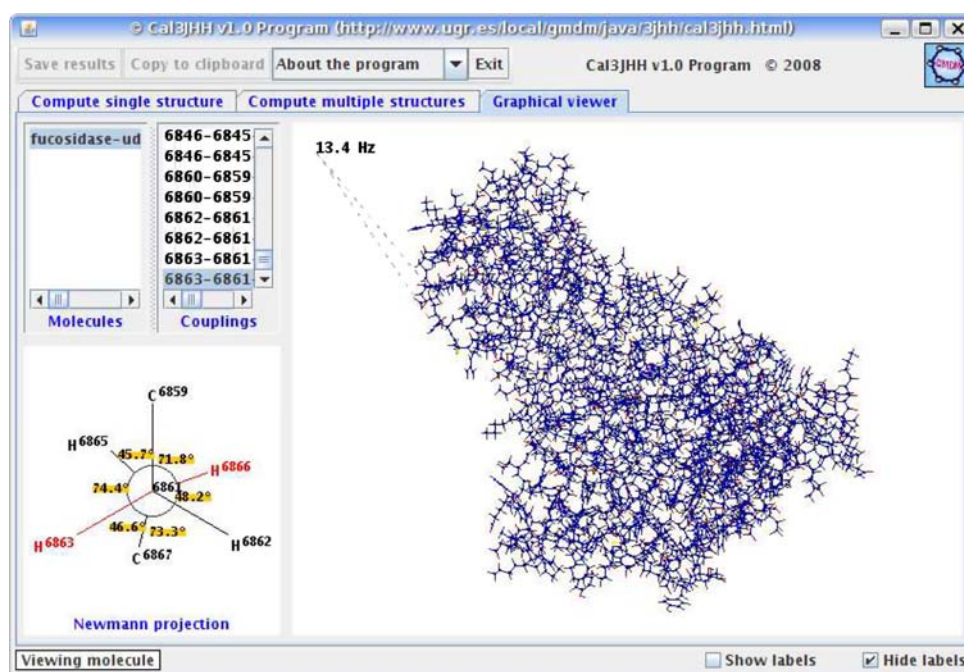
has 6872 atoms, and only from the PDB information the program automatically calculated the connectivity, and then recognized and computed 2385 couplings constants, and 214 averaged values.[14]

Also, the calculation can be restricted to a determined region of the molecule by previously defining the corresponding specific connectivity, and using the "From CONECT record" option. For instance, in the case that the PDB file has information about a ligand such as a carbohydrate, this option could be useful.

The second additional example stresses how this program provides substantial help in the conformational analysis. We recently reported a conformational study of thiopyranoses, where CAL3JHH has been successfully used [21]. The determination of the vicinal coupling constants, for these structures, is essential to elucidate the ring conformation, the anomeric configuration α/β, and the influence of the rotamers of the hydroxyl groups.

As an example, Table 1 shows the calculated data computed by the CAL3JHH program for the 6-deoxy-3-O-methyl-5-thio-α-L-idopyranose, and the available experimental values [22]. The experimental evidence for L-ido-thiopiranose configuration indicates the existence of a mixture of ring conformations in solution: $^4C_1$, $^1C_4$, and ^2S_S [23].

Table 1 clearly shows how the averaged values are in a better agreement than the values for the different conformations ($^4C_1$, $^1C_4$ or ^2S_S). This fact underlines the importance to take into account the rotamers populations,

Table 1 Summary of results for the $^3J_{H,H}$ (Hz) of the 6-deoxy-3-O-methyl-5-thio-α-L-idopyranose, made with the CAL3JHH PROGRAM

Conformation	%	$J_{1,2}$	$J_{2,3}$	$J_{3,4}$	$J_{4,5}$
$^4C_1$	84.57	8.9	9.3	9.3	4.4
$^1C_4$	10.79	2.8	3.5	3.7	1.6
^2S_S	4.64	6.4	10.1	5.0	3.5
Averaged values		8.1	8.7	8.5	4.1
Averaged values for the 35 rotamers		7.6	8.0	7.9	3.8
Experimental data [22]		8.0–8.3	8.4	8.2–8.6	4.6

and the utility of the program to automatically carry out the calculations.

Conclusions

The CAL3JHH program for the vicinal $^3J_{H,H}$ coupling constants, here presented, shows its utility in the conformational analysis, because it accounts for the averaged $^3J_{H,H}$ values of the Boltzmann populations. Besides the user-friendly design, the introduction of a graphical viewer converts the CAL3JHH into a practical tool for the conformational analysis of a single or multiple molecules. Furthermore, its development in a JAVA programming language together with the free web access facilitates information exchange throughout the scientific community.

---

[14] This process, performed in a notebook with the Intel Core 2 Duo processor (1,5 GHz, 667 MHz, 2 MB), lasted six seconds.

References

1. Eberstadt M, Gemmecker G, Mierke DF, Kessler H (1995) Angew Chem Int Ed 34:1671
2. Karplus M (1959) J Chem Phys 30:11
3. Duus JØ, Gotfredsen CH, Bock K (2000) Chem Rev 100:4589
4. Altona C, Francke R, De Haan R, Ippel JH, Daalmans GJ, Hoekzema AJAW, Van Wijk J (1994) Magn Reson Chem 32:670
5. Hassnoot CAG , De Leeuw FAAM, Altona C (1980) Tetrahedron 36:2783
6. Huggins ML (1953) J Am Chem Soc 75:4123
7. Abraham RJ, Fisher J, Lofitus P (1997) Introduction to NMR spectroscopy. Wiley
8. (a) Osawa E, Jaime C, 3JHH Program, Department of Chemistry, Faculty of Science, Hokkaido University, Sapporo 060, Japan; (b) Keisuke I, Osawa E, 3JHH2 Program, Department of Chemistry, Faculty of Science, Hokkaido University, Sapporo 060, Japan; (c) Petillo PA, Osawa E, Jaime C, 3JHHPC Program, Department of Chemistry, University of New Hampshire, Durham, New Hampshire 03824 and Department of Chemistry, Faculty of Science, Hokkaido University, Sapporo 060, Japan
9. Cerda-García-Rojas CM, Zepeda LG, Joseph-Nathan PA (1990) Tetrahedron Comput Method 3:113
10. Balacco G (1996) J Chem Inf Comput Sci 36:885
11. Cramer CJ, Truhlar DG (1993) J Am Chem Soc 115:5745
12. Tinoco I, Jr., Saner K, Wang JC (1978) Physical chemistry: principles and applications in biological sciences. Prentice Hall, New Jersey
13. Melchor S, Dobado JA (2004) J Chem Inf Comput Sci 44:1639. CONTUB program: available on http://www.ugr.es/~gmdm/java/contub/contub.htm
14. http://www.ugr.es/local/gmdm/java/3jhh/cal3jhh.html
15. http://java.sun.com
16. http://www.sun.com
17. http://www.rcsb.org/pdb
18. Gaussian 98, revision A.7 (1998) Gaussian, Inc., Pittsburgh, PA
19. SDBS Database, National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, Japan, http://www.aist.go.jp/RIODB/SDBS/menu-e.hmtl
20. Sulzenbacher G, Bignon C, Nishimura T, Tarling CA, Withers SG, Henrissat B, Bourne Y (2004) J Biol Chem 279:13119
21. Aguirre-Valderrama A, Dobado JA (2006) J Carbohydr Chem 25:557
22. Isac-García J, Calvo-Flores FG, Hernández-Mateo F, Santoyo-González F (1999) Chem Eur J 5:1512
23. Hughes NA, Munkombwe NM, Todhunter ND (1991) Carbohydr Res 216:119