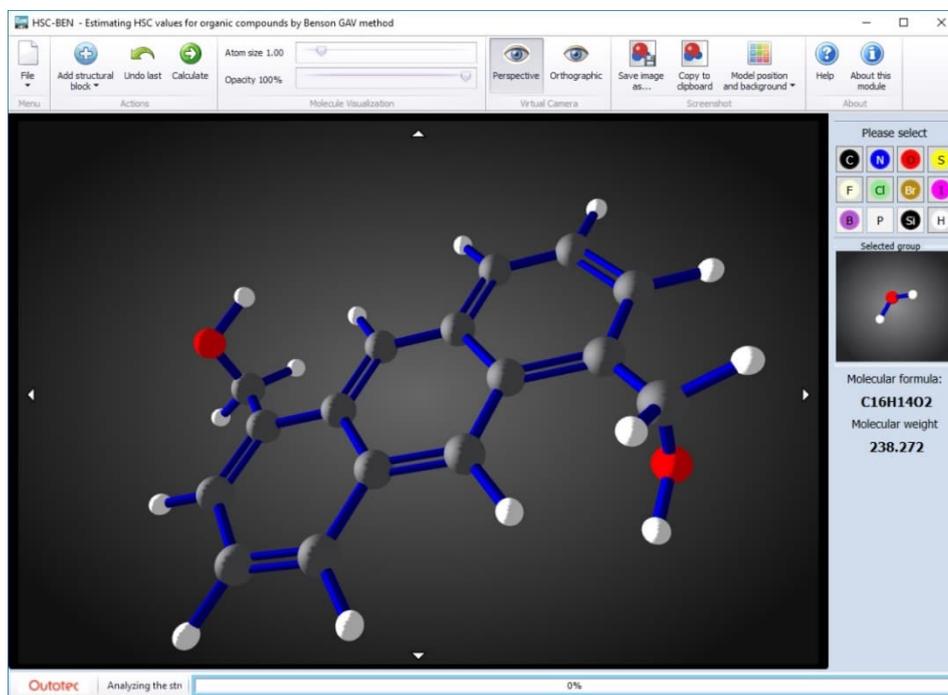


23. Benson Module



23.1. Preface

In spite of the vast extent of the comprehensive database included in HSC, which is the largest available in the commercial world, only a small fraction of all known organic compounds have ever undergone laboratory measurements to find out their basic thermodynamic properties. Therefore we decided to include in HSC Chemistry a new module, which in most cases calculates reliable estimates of the standard enthalpy of formation in normal conditions, entropy and specific heat capacity for many groups of organic compounds using the Benson Group Additivity theory and method.

23.2. Introduction

The number of characterized organic species in the Chemical Abstracts Registry now exceeds 158 million and new ones are being added at a rate of several millions every year. On the other hand, the number of species for which basic thermodynamic quantities have been experimentally measured can be estimated to amount to several thousand at most and yearly additions to this list in the tens or hundreds. This fast growing gap between the extent of available experimental data and the need for information has influenced the development of various empirical and semi-empirical methods to estimate the required data. One of the most widely used and valued of these methods was originated by Sidney W. Benson and later extended and enhanced by many other researchers, the most notable of whom are E. Domalski, E. Hearing, and N. Cohen. As the name of this module suggests, it is based on this group additivity theory and method. Such kinds of theories and methods are based on the idea that any organic molecule can be split into a limited number of different groups, which consist of a central atom and a variable number of ligands, so that the basic thermodynamic quantities may be estimated with sufficient reliability by adding together the contributions made by each constituent group and taking into account some correction terms where needed.

23.3. Benson Method

The most often referred to and most widely used group additivity method¹ in chemical publications for estimating thermodynamic properties is named after S. W. Benson, who was its most prominent developer. Since its original publication, this method has been extended to other phases and its application area widened several times²⁻⁵. One of the biggest advantages of this method is that it practically never results in major anomalous deviations among all members of compound groups belonging to its application area; the error margin stays very reliably within reported limits. Recent attempts to develop this method to a level where tabulated values for C_p would not be needed for a large number of temperature ranges have not been successful and no generally valid way to calculate C_p as a function of temperature has been found. The greatest inconvenience in adapting Benson method for daily work is the large number of the various atom groups used. To utilize the Benson method smoothly without the help of a computer, the user would have to memorize a list of hundreds of groups in order to see immediately how to split the required molecule into valid groups, for which the additivity values can be found in Benson tables. In addition to this, the method is not purely additive, as there are a couple of correction terms and tens of extra terms for ring strain corrections and to account for the interactions of some pairs of groups, especially *cis*- isomers of alkenes as well as *ortho*- and *meta*- isomers of benzene derivatives.

The core formulae of the Benson method are shown in three equations:

$$S^0(298,15\text{ K}) = \sum_j n_j \Delta_S - R \ln \sigma + \ln \eta \quad (1)$$

$$\Delta H_f^0(298,15\text{ K}) = \sum_j n_j \Delta_H \quad (2)$$

$$C_p^0(298,15\text{ K}) = \sum_j n_j \Delta_{C_p} \quad (3)$$

where n is the number of groups of type j , Δ the group contribution, σ the molecule symmetry number, and η the number of possible optical isomers. R is the general gas constant according to the kinetic gas theory. Besides finding the correct combination of Benson groups and looking up the table values, a fair amount of additional work is caused by the need to calculate the symmetry number of the molecule and the number of optical isomers, which slows down the application of this method in daily routine work. This also makes it more difficult to implement the method in a computer program. On the other hand, a fair amount of the most common cases requiring these terms can be covered by tabulating values inside the program, so it is not compulsory to build a lot of intelligence into the code in order to obtain reasonable values for these additional terms.

The correction term needed to calculate entropy can be derived using statistical mechanics that define entropy as $R \ln W$, where W is the number of distinguishable configurations of the molecule⁶. As all rotational, oscillating and other conceivable movements on molecular level are quantified, only a finite number of configurations defined by these exist. For practical calculations, it is feasible to divide the correction term σ into two factors: $\sigma = \sigma_{int} \sigma_{ext}$ separating it into permutations of configurations of possible intramolecular groups and those of the whole molecule.

23.3.1. Starting the program

As the module starts, most of the main window is taken up by a dark area, where construction of the molecular model will take place. The right side vertical bar of this area contains most of the buttons and controls needed to build a model. Above the construction area, there is a ribbon type control center containing the buttons most often needed to control the work flow with this module and the visual appearance of the model. At the bottom of the window there is an indicator telling the user how the estimation calculations are proceeding.

In the middle of the right side there is a small window, where the user can see an atom or atom group rotating. This window always depicts the group the user selected to be added to the model next. The default group shown at the start of the program is a methane molecule, which is a common starting point for many organic species. Above this window there are ten clickable buttons, behind which the user will find all the ready-made molecule fragments and select the most suitable starting group for the intended molecule. As it is not possible to construct new ring structures with this program in its current phase of development, tens of the most commonly needed cyclic molecules are found under the 'Add structural block' drop-down menu. Most common functional groups handled by the Benson method can also be selected under the 'Functional group' submenu. At the moment it is not allowed to start construction with a halogen or hydrogen atom, so these buttons become active immediately after addition of the first group to the model.

23.3.2. Molecule construction

When a suitable group to start the construction has been found, it can be placed in the center of the construction area by left-clicking the dark area. Now all free bonds (valences) are filled with hydrogen atoms. The current version of this module only allows you to build one model at a time, so clicking the background does nothing once the first group has been inserted into the model. After this new groups may be added to the model only by replacing existing hydrogen or halogen atoms. An atom is replaced by a new group by clicking the appropriate atom in the model. You can continue by adding more of the same skeletal groups or by changing the selected group. All the single atom substituents supported by this program have buttons of their own on the right side of the window. There are also some buttons for the simplest multi-atom substituents. When constructing a more crowded structure, please take care and observe the atoms on the edges of the last added group. If you see some atoms overlapping, we recommend undoing some of the last additions and trying a different order or placement, because overlapping atoms are sometimes interpreted by the estimation program as two atoms with bonds to the same carbon, which causes erroneous identification of the atom groups and probably incorrect estimates.

23.4. HSC estimates and Own database

When the molecule structure is completed, click 'Calculate' to start the analysis of structure and estimation calculations. These will take a moment. During the analysis another button will appear for aborting the current work. If you see that the calculation is taking too long or notice at this stage that something is missing from the model, you can interrupt analysis and make changes as you wish. When the analysis and calculations are ready, the application will next show you all the species found in the HSC and Own databases corresponding to the gross molecular formula of the entered molecule with a temperature range somewhere around room temperature.

At this stage the program will not filter the species to adapt to the structure, if there are not more than twenty choices, so in some cases you will see a lot of unnecessary rows. If the list becomes longer, there will appear an "Apply filter" button and text box on top of the list. So, if you for instance know having entered a substituted hexane, you just need to type "hex" to the text box and click the button. Now the program will filter out from the table every row not containing word "hex" and You will get more easily manageable listing. You need to be careful not to enter too long filter criterion, but rather use a short one. Otherwise You may accidentally filter out also the correct reference species. At this stage the filtering can only be done once with one species list. In case of error You have to cancel the estimation and start again with "Calculate" button. If no species in either database corresponds to the formula you entered, this phase will be skipped. If just one suitable species is found, its data values will be forwarded to the next phase without any notification. If a list is shown, you should select the most relevant species for comparison with the estimated values for each phase, if available. When you have marked the best available comparison set, or nothing in some cases, click the 'Ready!' button at the top of this window.

The program will next show the estimates and display them paired with the corresponding values extracted from the HSC database. Now you may select a set of estimates that you think valid to be added to your own database. Selection is done by clicking the row with the desired phase abbreviation. This will mark all estimates on this row as selected. If you next wish to save them, click the 'Add selected estimates to own database' button. This button also has a dropdown list of possible base records from the database, if available, so you can select a suitable starting point for the additional data needed. Please remember that this program calculates estimates only for three basic thermodynamic quantities. All other information, like name, CA Number, melting and boiling points etc. in the database entry will be taken from the HSC or the Own database records that the user chooses here, or the user will need to enter them if no suitable base record is found. Further editing will take place in the HSC standard database editor window, which will open next.

Please note that this program version only allows you to insert full sets of three values into the database. If even one of these values has been calculated using insufficient data, the application will not allow you to insert these uncertain values into the Own database. If you wish, it is of course possible to take notes or copies of the estimated values and insert them into your own database manually using the database editor of HSC Chemistry. The intention is for later releases of this module to include a more complete database of group additive values as they become available, so this limitation should be reduced during the coming years.

The estimation result window also offers a button to directly copy its contents to the Windows clipboard to be used in other applications, if required. This simplifies the task a little, as you do not need to select anything with the mouse.

Editing of records in the Own database is delegated to the Database module of HSC, which can be accessed from the application main menu.

23.5. Main menu

The 'File' menu allows you to save the structure you built for later use. The program will show a default name, but it is highly recommended to change it to something that better describes the species you built. Of course you also can load these files back to this application to continue working with them. The structure is always saved in two different formats, each with its pros and cons.

This program has its own format (*.ben8), which saves every group you have inserted into your model in correct order, except those deleted by the 'Undo' button. So the program should be able to reproduce exactly what you did and after reading the model from the file, it is possible for instance to undo one or more of the last additions to construct a close derivative of the former species. Unfortunately, this file format is not yet 100 per cent reliable. Sometimes you may notice empty bond ends or other errors in the resumed model.

For this reason, there is another possibility, the nearly standard Chemical Markup Language (*.cml) file format. It will save the completed model as one group, which cannot be reverse engineered. This format of course has the benefit of offering a way to migrate your models to third party applications. The generated file is a highly simplified one and only contains data of atom types with their coordinates and bonds. In addition to these, there are two groups of data that do not belong to the standard, but allow the model to be analyzed and estimates calculated again by this application. This also means that in the current stage of development, this module cannot accept new CML files saved by third party applications directly as models. The BEN module will read and render them as graphics, but probably no reliable estimations can be produced.

There is still another way to save models, as images. These are of no use in structure analysis, but enable you to add pictures of your models to various documents and publications. You can save the model directly from this module in a disk file or on the Windows clipboard. In both cases you can select the file format from four choices: Bitmap file (*.bmp), Portable Network Graphics (*.png), JPEG (*.jpg) and Graphics Interchange Format (*.gif). There is a program menu to select the format for the clipboard and for a disk file, the selection can be done using the File Save dialog. In both cases you can preselect a new background color for the model to replace the normal dark background. The background color change will be visible in the main window the moment the program pauses to save the image.

23.5.1. Other control buttons in the menu ribbon

Under the 'File' menu button there are four possible actions:

1. 'New' cancels the whole structure constructed so far and lets the user start over, after confirmation.
2. 'Read' lets the user read any previously saved structure in the working area.
3. 'Save' function is explained above.
4. 'Exit' will quit this program after user confirmation.

'Undo' will remove the last inserted group from the model in the structure window. The groups are chained so that you may remove every added group in reverse order one by one, if required.

23.6. Appearance of your molecule model

To create different views of your model you can choose between two types of virtual “cameras” through which you look at your molecule model. The camera types available are called ‘Perspective’ and ‘Orthographic’. To clarify these concepts, please take a look at the illustrations in **Fig. 1** and **Fig. 2** of the coronene molecule below:

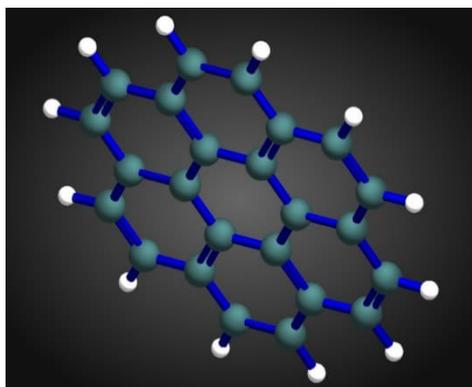


Fig. 1. Orthographic view

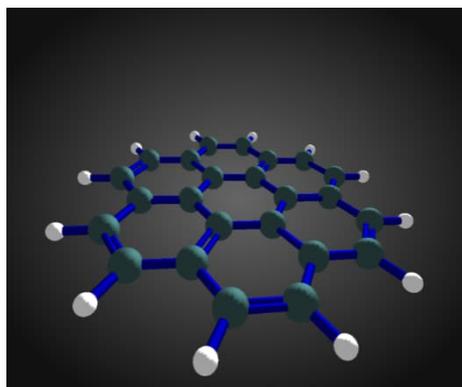


Fig. 2. Perspective view

One way to significantly change the appearance of your models is the choice between the traditional ball-and-stick model, which represents the bonds between atoms more clearly, and the space filling model, which probably gives a slightly more realistic impression of the molecule electron cloud. Changing between these models can be done with a slider control anytime during model construction. The opacity of space filling model atoms can be varied from 0 to 100 per cent with a slider control. The sphere size of the atoms can be varied from half the size normally used in the ball-and-stick model by default, up to five times this size. The ‘standard’ settings are 100 % for the opacity space-filling and 1.00 size for the common ball-and-stick model. For a space filling model, a value of 3.60 is a good starting point. Both types of models can of course be saved as images if needed.

The next pair of images demonstrates the differences between these two ways of presenting exactly the same model of cyclooctane in the same orientation, in **Fig. 3** and **Fig. 4**:

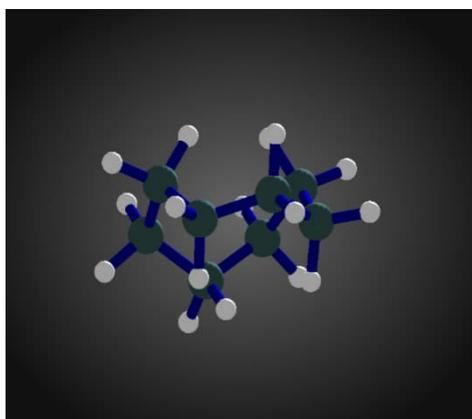


Fig. 3. Ball-and-stick model

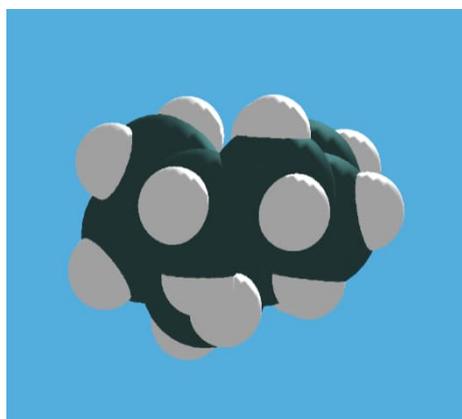


Fig. 4. Space filling, 100 per cent opacity, 3.6 x default atom size, custom background color

For both sliders, please note that the new value is only reflected in the appearance of the model after a small delay. Both sliders can also be moved by mouse wheel.

23.7. Limitations

The first version of this module was limited to handling organic compounds that contain carbon, hydrogen, oxygen, nitrogen, sulfur and halogens, due to the inherent limitations of the original Benson method. Many combinations of these are still missing, as the Benson method groups in available publications do not cover all possibilities at present. In the current version we have succeeded to extend the method into handling of some organic phosphorus (III) and (V) compounds, too. This extension partially relies on quantum chemical computations, as experimental data is very scarce. We are still continuously researching all possible sources of valid thermodynamic data to find more useful values and expand our group contribution database. The additional material will be available in updates and future releases of this program.

In the current version of this module it is not yet possible to control which position of the added substituent will be bonded to replace the clicked (hydrogen) atom. One example: If 1-methylnaphthalene is needed, the construction should be started with naphthalene itself. The methyl group can be then added to either the 1 or 2 position, but if the methyl is introduced first, the user can no longer control the position in which the substituent will end up and 2-methylnaphthalene will be drawn by default. This means that a lot of simpler molecules may be constructed with a greater or lesser degree of preplanning, but anything containing many groups and rings might be beyond the capabilities of this version.

We hope the users will understand that this is just the second published main version of this module and that the research is still intensive. Our goals and expectations of future capabilities and features of this module are far higher than the current incarnation of the BEN module.

NB! We would not encourage or recommend trying to use the BEN module for inorganic compounds as they are beyond the capabilities of Benson method itself.

Any feedback, bug reports and development suggestions will be greatly appreciated.

23.8. Simple Example

The aim is to estimate the thermochemical properties of 4-methylaniline (*4-methylbenzeneamine*).

The construction of the molecule can be started with the aromatic ring found in the Structural blocks (**Fig. 5**).

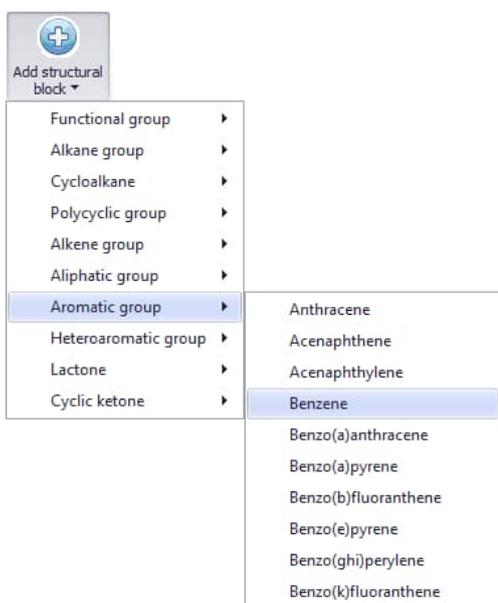


Fig. 5. Aromatic skeletal groups.

After selecting the *Benzene* structure, click the workspace to insert the first component of the structure (**Fig. 6**).

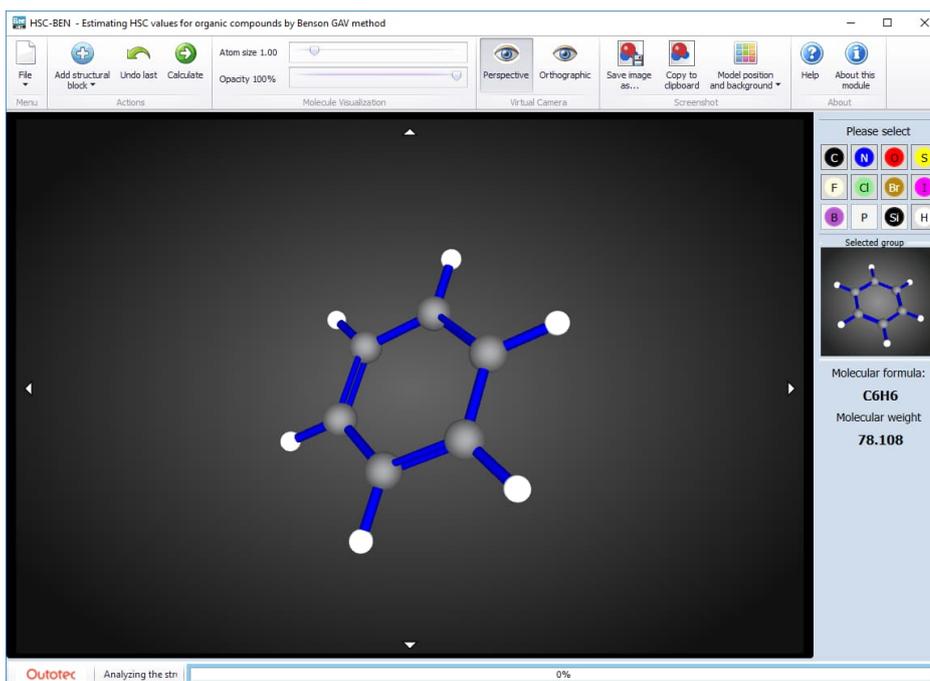


Fig. 6. Benzene ring.

Next, to complete the structure, the molecule requires a methyl group and an amine group which can be found from the single atom buttons (**Fig. 7** and **Fig. 8**).

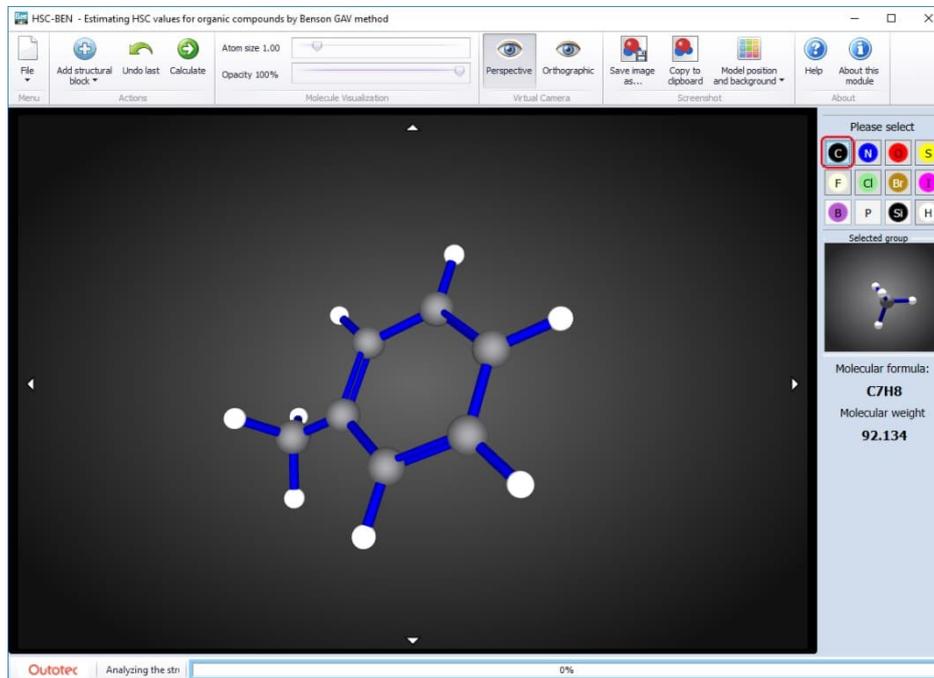


Fig. 7. Methyl group added.

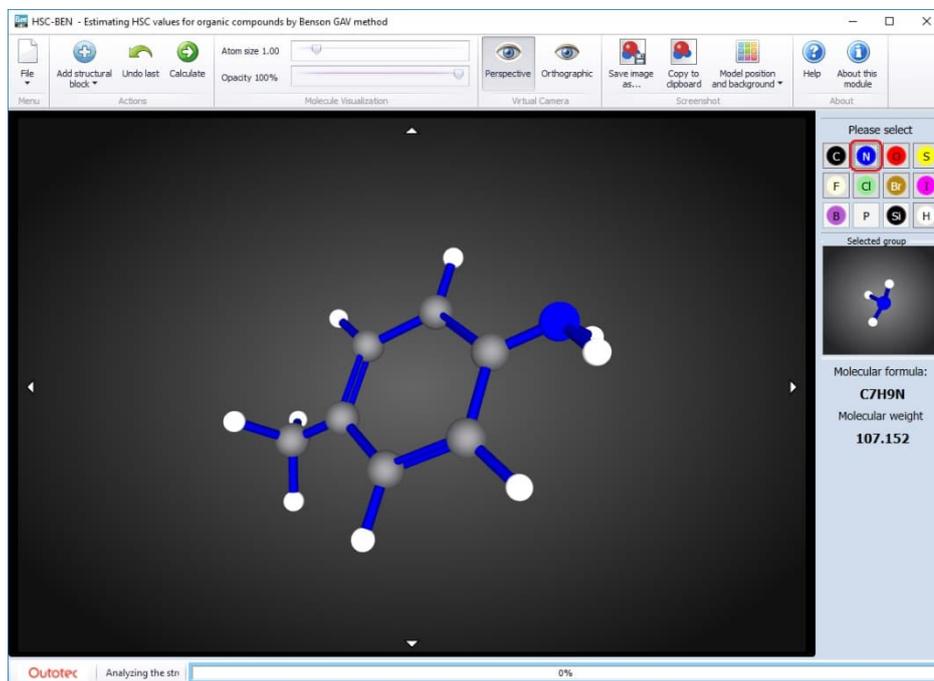


Fig. 8. Amine group added.

Once the molecule is ready, click "Calculate" from the top menu bar. From the species list that appears, select liquid 4-Methylaniline as a comparison reference for the estimated results (**Fig. 9**).

<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	C7H9N(4MAI)	CH3(C6H4)NH2		298.15	300	4-Methylbenzenamine	4-Methylaniline; p-Toluidine
--------------------------	-------------------------------------	--------------------------	-------------	--------------	--	--------	-----	---------------------	------------------------------

Fig. 9. Check HSC database species for the result.

The calculated estimates are presented in **Fig. 10**.

Estimation Results in joules

Formula Common name
 Chemical name

Phase,Source/Propert	ΔH_f^0 in kJ/mol	S^0 in J/mol*K	C_p^0 in J/mol*K
Gas, est.	54.57	349.18	130.34
Gas, db.			
Liquid, est.	-5.31	226.56	214.91
Liquid, db.	-5.28	226.56	214.91
Solid, est.	-18.57	174.62	163.71
Solid, db.			

Save to own database Copy to clipboard Joules Close

Fig. 10. Estimated results for 4-Methylaniline.

The estimated values can be saved to the Own database by selecting the correct row of estimates and clicking the "Save to own database" button. You can also choose to use one of the previously selected comparison species as a base reference for the new species. The base reference species lets you copy basic data (chemical name, common name, etc.) to the new species (**Fig. 11**).

Estimation Results in joules

Formula Common name
 Chemical name

Phase,Source/Propert	ΔH_f^0 in kJ/mol	S^0 in J/mol*K	C_p^0 in J/mol*K
Gas, est.	54.57	349.18	130.34
Gas, db.			
Liquid, est.	-5.31	226.56	214.91
Liquid, db.	-5.28	226.56	214.91
Solid, est.	-18.57	174.62	163.71
Solid, db.			

Save to own database Copy to clipboard Joules Close

No base reference

Fig. 11. Saving estimated values for the solid phase and using a base reference.

Finally, the data of the new species is presented in the HSC Database editor (**Fig. 12**), in which you can edit any of the values. Please note that you must click the "Save" button in the database editor to close the edit mode and save the Own database file, for the addition of the new species to take effect.

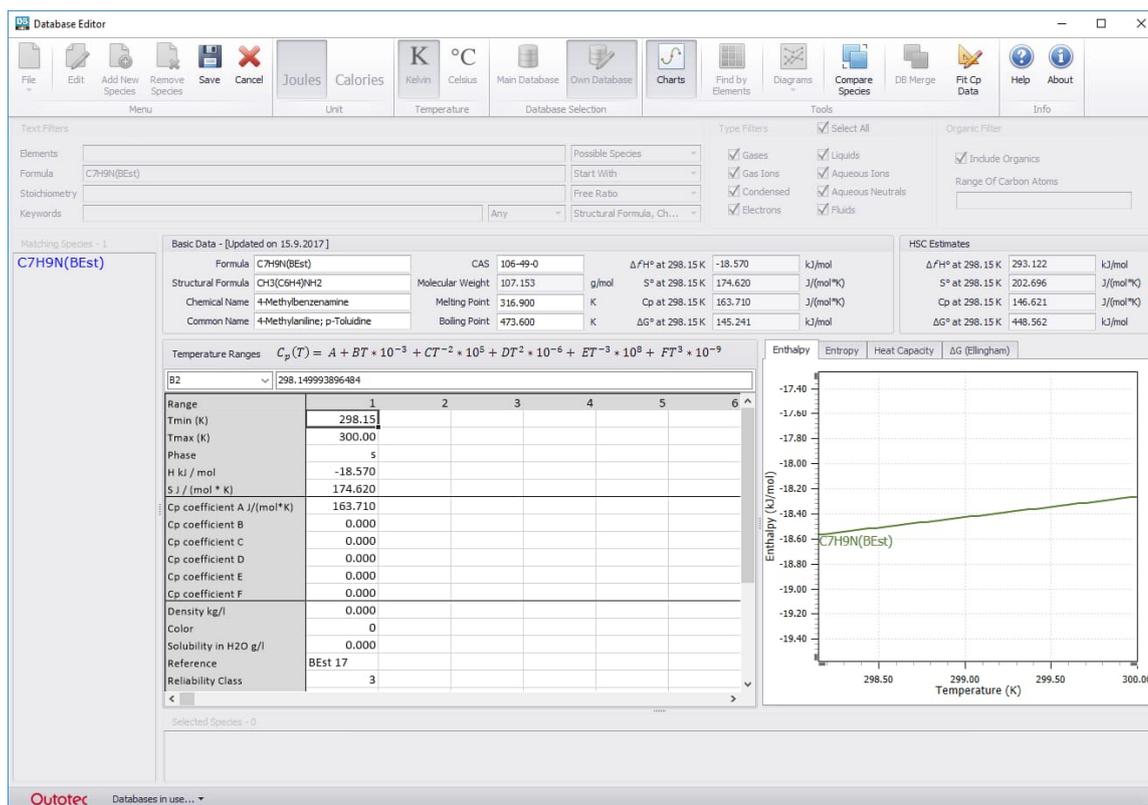


Fig. 12. Saving estimated values to the Own database.

23.9. Using computational chemistry in extending Benson method

Until recent years all researchers widening the application area of Benson method have kept a strict principle of using only experimental data from laboratories in deriving new group values. As the flow of new species gets year by year more overwhelming compared to the modest stream of new laboratory measurements of the fundamental thermodynamic values, there has for decades been waxing interest in substituting or complementary methods to produce new data. The most experienced masters of experimental methods are getting older and one by one retiring from scientific arena and the younger generations seek more modern solutions.

The next science conquering the area formerly ruled by laboratory experiments in thermochemistry, reaction mechanisms and many more sections of chemistry, seems to be computational chemistry. This means that all the properties of a molecule or a bulk of some species are calculated starting from wave functions of atomic and molecular orbitals or related entities, like in Density Functional Theory (DTF)⁷.

This requires usually first optimizing the molecular structure and conformation to its minimum total energy. In ab initio methods this calculation is based on foundations of the famous Schroedinger wave equation. In a quite normal case this may on a deeper level of theory take tens of iterations each containing billions of integral function expression solving calculations. Only in recent decades have these computational methods and computing power of the hardware reached a level, where decent results can be obtained in reasonable time and with reasonable investment into computing resources. In this case a reasonable time means sometimes several weeks for one calculation using a decent computing cluster with formidable resources. It is no wonder that some of the most prominent computational chemists have been nominated as Nobel laureates.

In DFT the calculations also aim to minimize the total energy of the system under work, but now the principle is based on the assumption that the energy of a system is only dependent of its electron density. This can even be mathematically proven. Then we only need to find a functional accurately describing this electron density at every point of the system. The final functional will never be found, but it is possible to closer and closer simulate it with better functionals, up to the level, where results from these calculations are accurate enough to give scientifically noticeable outcomes. DFT has already produced new inventions in many areas of chemistry, like superatoms and better understanding of catalytic processes. This has further led to better catalysts for industrial work.

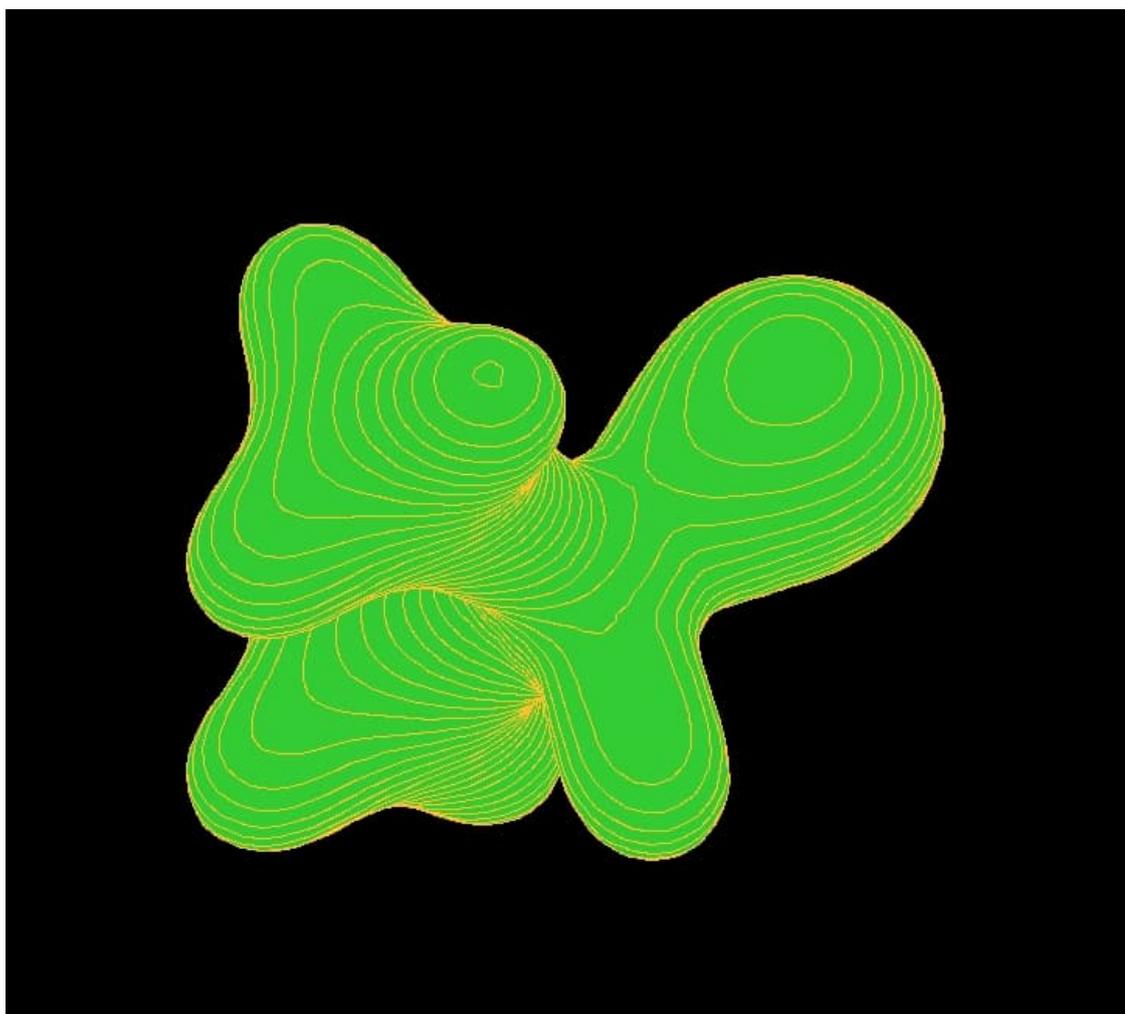


Fig 13. One of the molecular orbitals of dimethylphosphine as visualized by Molden program from the output of Gaussian09 application.

The prices and availability of powerful computing equipment and development of versatile and efficient software like Gaussian and MolPro have induced an exponentially growing flow of publications and articles containing new thermochemical data, which are often claimed and sometimes even proved to match or exceed the laboratory values in accuracy.

The pace of this development is rather getting faster than slower and every year we see larger and larger molecular systems successfully taken under calculations.

It is our intention in the next phases of development of this program module to further loosen the principle of only using experimental data in deriving new Benson group increment values and start using additions obtained by methods of computational chemistry. This will enable us to extend the application area of this module into organic boron, phosphorus and silicon compounds. In the results table values obtained by utilizing quantum chemically computed numbers are shown in blue color.

23.10. References

1. Benson, S. W.: Thermochemical Kinetics, Second Edition J. Wiley & Sons, New York, 1976, chapter 2.
2. Eigenmann, H. K., D. M. Golden, and S. W. Benson: J. Phys. Chem. 77: 1687 (1983).
3. O'Neal, B. A., H. E., and S. W. Benson: J. Chem. Eng. Data, 15: 266 (1970).
4. Shaw, R.: J. Phys. Chem., 75: 4047 (1971).
5. Stein, S. E., D. M. Golden, and S. W. Benson: J. Phys. Chem., 81: 314 (1977).
6. Denbigh, K.: The Principles of Chemical Equilibrium, Fourth Edition, Cambridge University Press, chapter 14.1, 1981.
7. <http://blogs.cimav.edu.mx/daniel.glossman/data/files/Libros/Computational%20Chemistry%20and%20Molecular%20Modeling%20-%20Principles%20and%20Applications%20.pdf>

The group contributions database used by this application internally was compiled from numbers published in:

- 1 Domalski, E. S. and Hearing, E. D., J. Phys. Chem. Ref. Data, 22, (4), 1992
- 2 Khalifa, A., Ferrari, M., Fournet, R., Sirjean, B., Verdier, L., and Claude, P. A., J. Phys. Chem. A 119, 2015
- 3 Dorofeeva, O. V., Ryzhova, O. N., and Moiseeva, N. F., Russ. J. Phys. Chem. A 82, (6), 2008
- 4 Vuori, H.T., Rautiainen, J.M., Kolehmainen, E.T., and Tuononen, H.M., Benson Group Additivity Values of Phosphines and Phosphine Oxides: Fast and Accurate Computational Thermochemistry of Organophosphorus Species, J. Comput. Chem. 2019, 40, 572-580