
# BIOTERRA


**Code NURC : 882 category „C"**
**ISSN:1454 - 816X**
**The journal is indexed in the following International Databases:**
**ReportLinker.com, SCOPUS, Google Scholar, INDEX COPERNICUS, CABI, EBSCO.**

## BULLETIN OF SCIENTIFIC INFORMATION

NR. 37 JANUARY-JUNE 2019

(twice a year publication)

**MUSTANG Publishing House**
Recognized by
National Council for Scientific Research in Higher Education (NURC), cod 181
mustang.expres@gmail.com

**Technical Editor:**
**BETERINGHE ADRIAN**
**adrian.beteringhe@gmail.com**

**Office:**
Nicolae Şuţu Street, no. 11 - 15
Central building, second floor, room 9,
sector 1, Bucharest, Romania
http://www.bioterra.ro/bulletin.php

Publication Recognized by NURC category "C" code NURC : 882
The journal is indexed in the following International Databases:
ReportLinker.com, SCOPUS, Google Scholar, INDEX COPERNICUS, CABI, EBSCO.
ISSN 1454 - 816X

**MUSTANG Publishing House**
Recognized by
National Council for Scientific Research in Higher Education (NURC), cod 181
mustang.expres@gmail.com

**Technical Editor:**
**BETERINGHE ADRIAN**
**adrian.beteringhe@gmail.com**

**Office:**
Nicolae Şuţu Street, no. 11 - 15
Central building, second floor, room 9,
sector 1, Bucharest, Romania
http://www.bioterra.ro/bulletin.php

Publication Recognized by NURC category "C" code NURC : 882
The journal is indexed in the following International Databases:
ReportLinker.com, SCOPUS, Google Scholar, INDEX COPERNICUS, CABI, EBSCO.
ISSN 1454 - 816X


***Rector's Allocution***

*We have the special pleasure to let you know that the Review of our University, „Bulletin of Scientific Information", having ten years of consecutive issue, it achieved the recognition of the National Council for Scientific Research in Higher Education (NURC), being comprised in the category „National Reviews — C Category".*

*So, Bioterra University review „Bulletin of Scientific Information" works as a real platform for the information and exhibition of the most recent and valuable research in the agricultural field and connected sciences (food industry, agro-tourism, ecology, environment protection, agricultural economics etc).*

*This way, I express my gratitude to the contributors to our science magazine, to the authoritative academic and universitary personalities of whose studies are found in the selection done by the scientific board of our magazine with whom we have strong relations of partnerships in the development of jointed research projects.*

*I wish to our scientific science magazine many and consistent issues.*

**Prof. Floarea Nicolae, PhD**

**Rector of Bioterra University Bucharest**

## *Editorial Board's Allocution*

*„Bulletin of Scientifc Information" was published at the initiative of several young researchers with the direct support of Bioterra University Board, having the first edition in 1998.*

*Years passed and this magazine has enriched continuously its scientific and didactic dowry becoming slowly but surely a veritable platform for academic information.*

*In 2008, this science magazine turned into a new more dinamic and attractive pattern, being published in special grafic features (full-color) and fully in English language. Also, since 2014, our science magazine benefits of a modern website:* www.bsi.bioterra.ro.

*Every year the editorial team has increased the number of members; nowadays it brings together numerous personalities of the scientific and academic world from different foreign countries, thus being a guarantor of a high scientific level.*

*Thanks to all our readers and collaborators that through their suggestions, criticisms and feedback contribute to the improving of our science magazine quality.*

**Prof. Petculescu Nicole Livia, PhD**
**Vice Rector of International Relations**

**Bulletin of Scientific Information**
2019; 37
Published online JUNE 2019 (http://bsi.bioterra.ro)

# SUMMARY



Erata*

Dintr-o eroare materială în versiunea tipărită a acestui număr a apărut ca autor domnul Lect. Dr. Colang George la articolul nr. 3. Necerem scuze pentru inconveniente, dumnealui fiind scos ca autor în versiunea online.

**Bulletin of Scientific Information**
2019; 37
Published online JUNE 2019 (http://bsi.bioterra.ro)

# A new molecular descriptor for modelling hydrophobic-hydrophilic balance for some blocked tripeptides

**BETERINGHE Adrian[1], NICOLAE Marian[1], PETCULESCU Nicole-Livia[1]**

[1]Bioterra University of Bucharest, Romania, Faculty for Control and Expertise of Food, 013722, Bucharest, Garlei Street

**To cite this article:**
Beteringhe Adrian, Nicolae Marian, Petculescu Nicole-Livia, A new molecular descriptor for modelling hydrophobic-hydrophilic balance for some blocked tripeptides. *Bulletin of Scientific Information*, No. 37, 2019, pp. 7-10.

**Abstract:** In this paper, a new parameter HLB (hydrophobic-hydrophilic balance) was obtained using various structural parameters calculated using the specific software. The HLB values were correlated involving linear regression, with the values determined experimentally in the specialized literature (logP) in case of 10 blocked tripeptides, obtaining a statistical model that is validated by the characteristic parameters (R, SD, F, R (CV)). The overwhelming results determine us to use the statistical model obtained for other classes of bioactive compounds.

**Keywords:** Hydrophobic-hydrophilic balance, Tripeptides, logP, amino acids, bioactive

## 1. Introduction

All living systems have as their central area the peptides that in the specialized literature will be called the drugs of the future [1]. For the design of bioactive peptides and implicitly of peptide drugs, an important instrument is the hydrophobic-hydrophilic balance (HLB) or the partition coefficient 1-octanol - water (logP) [2-3].

Amino acids are the most important components of peptides and for the design of bio-active peptides hydrophobic parameters are of essential importance [4].

## 2. Material and methods

### *2.1. Data set*

The hydrophobic parameter (logP) and the structures of the ten blocked tripeptides (Table 1) were collected from specialized literature [3, 6, 7].

**Table 1.** Structure of the 10 blocked tripeptides involved in the study, the experimental value for logP and the amino acids components

| No | Peptide structures | Exp. logP | Amino acid structures |
|---|---|---|---|
| 1 | VAA | -1.40 | Val (V) |
| 2 | VAV | -0.67 | Ala (A) |
| 3 | VIG | -0.45 | Ile (I) |
| 4 | ALV | -0.14 | Gly (G) |
| 5 | VFA | 0.06 | Leu (L) |
| 6 | AVI | -0.20 | Phe (F) |
| 7 | IFA | 0.52 | |
| 8 | GAV | -1.56 | |
| 9 | AGF | -0.71 | |
| 10 | IAV | -0.21 | |

## 3. Results and Discussion

In this paper a new relation for the calculation of the hydrophobic-hydrophilic balance (HLB) is presented (equation 1). It should be mentioned that the HLB parameter is just as important as the logP parameter for determining those hydrophobic-hydrophilic portions of any bioactive molecule.

$$HLB = \left(\frac{1}{SA}\right)^{\frac{1}{VDW}} + (M - D) + SF - AlogP \quad (1)$$

where: SA is the surface area, VDW is 1.4 selling der Waals interactions, M represents a statistical parameter whose value is presented in Table 2 and depends on the molar refractivity (MR) determined for each tripeptide, D represents the moment of dipole, $SF_{f(ST)}$ is a statistical factor which depends on the surface tension ST according to equation (2) and AlogP is the Ghose-Crippen octanol-water partition coefficient.

$$SF = INT\left[\left(\frac{ST}{23}\right) - 0.5\right] \quad (2)$$

**Table 2.** The value of the molar refractivity (MR) ($cm^3$), of the statistical parameter M for the 10 blocked tripeptides

| No | Peptide structures | MR ($cm^3$) | M | SF |
|---|---|---|---|---|
| 1 | VAA | 68.68 | -0.4 | 1 |
| 2 | VAV | 77.90 | 0.9 | 1 |
| 3 | VIG | 77.95 | 0.9 | 1 |
| 4 | ALV | 82.54 | 2.2 | 1 |
| 5 | VFA | 93.17 | 3.2 | 1 |
| 6 | AVI | 82.54 | 2.2 | 1 |
| 7 | IFA | 97.80 | 3.2 | 1 |
| 8 | GAV | 68.26 | -0.4 | 1 |
| 9 | AGF | 79.35 | 0.9 | 1 |
| 10 | IAV | 82.54 | 2.2 | 1 |

The SA, ST and MT values were calculated using the ACD / ChemSketch Freeware program [7]. Arguslab program [8] was used for the calculation of parameter D and AlogP is the Ghose-Crippen octanol-water partition coefficient, calculated using Dragon program [9]. Table 3 presents data obtained for the HLB construction presented in equation 1.

**Table 3.** The value of the parameters involved in the construction of HLB (equation 1)

| No | Peptide structures | Exp. logP | Surface Tension (ST) | Dipole Moment (D) | AlogP | SA | 1,4 vdW |
|---|---|---|---|---|---|---|---|
| 1 | VAA | -1.40 | 47.4 | 2.544 | -0.002 | 436.27 | 9.209 |
| 2 | VAV | -0.67 | 44.6 | 2.485 | 0.841 | 485.01 | 13.772 |
| 3 | VIG | -0.45 | 45.5 | 2.280 | 0.919 | 539.59 | 14.070 |
| 4 | ALV | -0.14 | 43.9 | 2.463 | 1.229 | 543.27 | 14.259 |
| 5 | VFA | 0.06 | 50.6 | 3.220 | 1.554 | 490.68 | 18.516 |
| 6 | AVI | -0.20 | 43.9 | 2.892 | 1.297 | 574.33 | 14.413 |
| 7 | IFA | 0.52 | 49.6 | 2.627 | 2.010 | 508.44 | 19.334 |
| 8 | GAV | -1.56 | 44.20 | 2.461 | -0.380 | 498.22 | 11.264 |
| 9 | AGF | -0.71 | 57.10 | 2.813 | 0.333 | 508.61 | 16.486 |
| 10 | IAV | -0.21 | 43.9 | 2.840 | 1.297 | 549.97 | 14.552 |

For the 10 structures presented in Table 1, the values of the HLB parameter were obtained using equation 1 which

will be presented in Table 4.

**Table 4.** The value of the HLB parameter (equation 1) obtained for the 10 structures presented in Table 1

| No | Peptide structures | Exp. logP | HLB (eq. 1) |
|---|---|---|---|
| 1 | VAA | -1.40 | -1.425 |
| 2 | VAV | -0.67 | -0.787 |
| 3 | VIG | -0.45 | -0.659 |
| 4 | ALV | -0.14 | 0.151 |
| 5 | VFA | 0.06 | 0.141 |
| 6 | AVI | -0.20 | -0.345 |
| 7 | IFA | 0.52 | 0.287 |
| 8 | GAV | -1.56 | -0.904 |
| 9 | AGF | -0.71 | -0.560 |
| 10 | IAV | -0.21 | -0.288 |

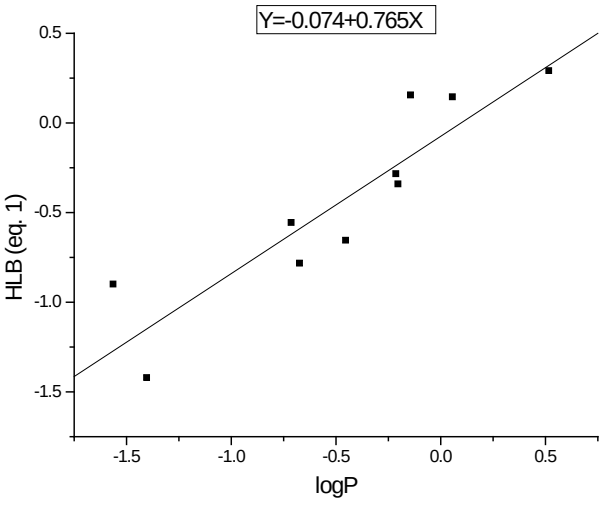



**Figure 1.** Graphical representation LogP vs. HLB (eq. 1)

Following the statistical analysis using the linear regression method, the statistical model presented by equation 3 is obtained, which is validated by the specific statistical parameters.

$$HLB\ (Y) = -0.074 + 0.765(\pm 0.126) \cdot X(logP) \quad (3)$$

R = 0.905 SD = 0.242 F = 36.631 R(CV) = 0.893

where R represents the correlation coefficient by linear regression, SD represents the standard deviation, F represents the Fisher test and R (CV) represents the cross-validation of the correlation coefficient R.

## 4. Conclusions

With the help of linear statistical analysis, a new parameter was calculated which represents the hydro-phobic-hydrophilic balance (HLB) in the case of blocked tripeptides. For the calculation of the HLB parameter, several structural parameters with important role in the structural interactions were used. HLB is very important in the design of bioactive peptides.

# 5. References

[1] S. Hellberg, M. Sjostrom, B. Skagerberg, S. Wold, *Peptide Quantitative Structure-Activity Relationships, a Multivariate Approach*, J. Med. Cherm., vol. 30, 1987, pp. 1126-1135.

[2] N. Sewald, H.D. Jakubd, *Peptides: chemistry and biology*, Wiley – Vch Verlag, 2002.

[3] J. Yin, *LogP Prediction for Blocked Tripeptides with Amino Acids Descriptors (HMLP) by Multiple Linear Regression and support Vector Regression*, Proc. Environ. Sci., vol. 8, 2011, pp. 173-178.

[4] T. Sotomatsu-Niwa, A. Ogino, *Evaluation of the hydrophobic parameters of the amino acid side chains of peptides and their application in QSAR and conformational studies*, J. Molec. Str. (Teochem), vol. 391, 1997, pp. 43-54.

[5] M. Akamatsu, T. Fujita, *Quantitative analysis of hydrophobicity of di- to pentapeptides having un-ionizable side chains with substituent and structural parameters*, J. Pharm. Sci.., vol. 81, 1992, pp. 164-174.

[6] M. Akamatsu, Y. Yoshida, H. Nakamura, et. al. *Hydrophobicity of di- and tripeptides having unionizable side chains and correlation with substituent and structural parameters*, Quant. Struct. Act. Relat., vol. 8, 1989, pp. 195-203.

[7] ACD/ChemSketch Freeware, Version v. 12.01 Advanced Chemistry Development, Inc., 110 Yonge, street 14th Floor, Toronto, Ontarion, Canada, http://www.acdlabs.com/.

[8] M. A. Thompson, ArgusLab, Version 4.0.1. Planaria Software LLC, Seattle, WA, 2004, http://www.arguslab.com/.

[9] R. Todeschini, V. Consonni, M. Pavan, A. Mauri, D. Ballabio, Dragon. Version 5.5. Talete srl, via V. Pisani, 13-20124, Milano, Italy, 2007, http://www.talete.mi.it/.

**Bulletin of Scientific Information**
2019; 37
Published online JUNE 2019 (http://bsi.bioterra.ro)

# Electronic archiving an efficient solution for all economic entities

**LĂCĂTUȘ Lorin-Nicu**

[1]Sistec NextDocs, Romania, Department of Archival Processing, Postal address1, 012352, Bucharest, Bucureștii Noi Boulevard

**To cite this article:**
Lăcătuș Lorin-Nicu, Electronic archiving an efficient solution for all economic entities. *Bulletin of Scientific Information*, No. 37, 2019, pp. 11-14.

**Abstract:** Electronic archiving has proven to be an efficient solution to classical archiving because through it you can archive a much larger quantity of documents, the process is faster and institutions prefer this type of archiving more and more.
For the practical realization of electronic archiving several stages are used but it has many more advantages over classical archiving, one of these advantages being the much larger quantity of archived documents or the faster finding of a document in the realized archive.

**Keywords:** electronic archiving, scan, indexing, storage, archival background

## 1. Introduction

All the institutions of a state are in a continuous dynamic regarding the daily activities and all these activities require documents [1]. For efficiency, all documents from the constitutive documents to the daily correspondence must be managed by constituting the so-called archive [2]. By definition the archive represents the totality of the documents created and / or held over time by any legal person, in the exercise of his activity or any natural person, during its existence [3].

The very large number of documents generated within a legal entity in conjunction with the development of information technologies has led to the emergence of the concept of electronic archiving by which the classical means of archiving documents have been replaced by modern procedures [4].

Electronic archiving of documents involves the following steps [5]:

- scanning / importing documents;

- indexing of documents;

- their storage on high capacity storage media and characterized by a very high reliability and their retrieval according to different criteria.

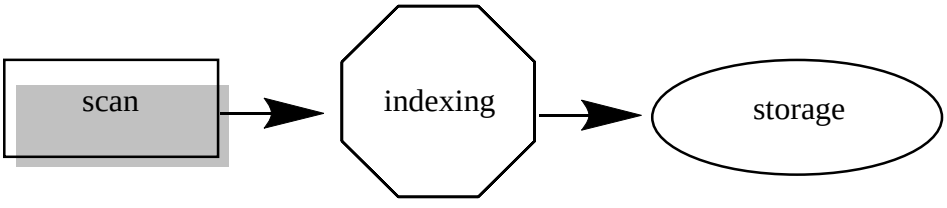



**Figure 1.** Stages in electronic archiving of documents

## 2. Results and conclusions

The electronic archive [6] can be defined as a multimer representing the electronic archiving system and all archived documents (metadata representing the context, content, structure of electronically archived documents and their administration over time) the binder between the two components representing the information of associated audits. The main element for the realization of the electronic archive is the electronic archiving program that allows easy retrieval of the documents stored in the archive according to a series of criteria involving the name of the document or keywords.

### 2.1. Electronic archiving directions

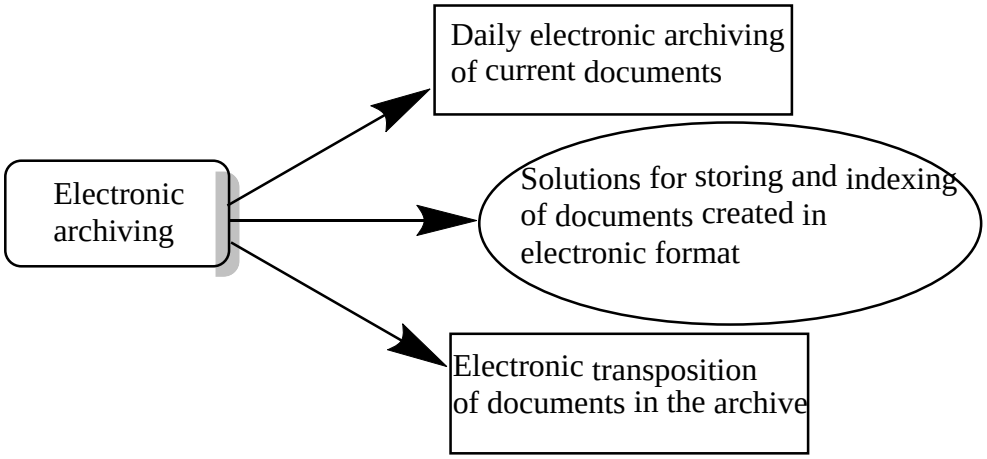



**Figure 2.** The utility of electronic archiving

### 2.2. Stages of the electronic archiving process

#### 2.2.1. Analysis of the archival background and establishing the scanning and indexing modalities.

In this phase, based on the visualization of the documents that are part of the archival background to be digitized, the formats for capturing (scanning) the documents from archival units, the name of the electronically archived logical entities, the name of the indexing and associated indexes templates, as well as the storage and archiving information in an electronic environment, under perfect security conditions.

#### 2.2.2. Preparing documents for scanning

If the archival fund to be digitized has not been processed archivist (constituted in archival units), it will be sorted, sorted and grouped according to the archival Nomenclature (if any). The documentation will be divided into broad categories, and within each category by years and types of documents.

In most cases, electronic archiving of documents follows the stage of archival processing, so that the records and digitization procedures will follow the specifics of the physical processing.

#### 2.2.3. Capture of documents

Documents are captured by scanning (scanning), duplex (if applicable), black and white or color. As a rule,

300 dpi resolution will be used for A4 formats, and for related documents (book type, description, registers, etc.) or larger formats (A3 and above) - higher resolutions between 300-600 dpi.

Depending on the status of the documents and their importance, the scan resolution is set. For historical documents, handwritten, color scanning formats (Jpeg), or grayscale (8-16-256) are used.

After verifying the scanned documents, they are assembled into logical entities, electronically archived, in final format (multipage TIFF format, or multipage PDF format).

#### 2.2.4. Restore the archive units ready for scanning

After capturing the documents, the files / entities prepared for the scan are restored to their original form (reconstruction of archival units).

If the procedure for electronic archiving of documents is carried out in conjunction with the previous procedures for physical processing, the scanned files are transmitted to the team of verification and binding of the archival units, continuing the archiving processing operations..

#### 2.2.5. Indexing information from scanned documents

Indexing information from scanned documents can be done by keywords, or content, within indexing templates per file (archive unit), or per document / archive logical entity:

• The templates and indexes of information are developed for each file (archive unit), or per document / archive logical entity;

• The information indexes (usually 5 indexes) are filled in as the documents are inserted in the electronic archive. The name, size and content of the indexes will be established together with the Beneficiary at the Analysis phase of the archival fund for digitization.

#### 2.2.6. Electronic archiving of digitized documents and associated indexes

Electronic archiving of documents and indexes is done on an existing hardware and software platform.

The images of the documents obtained by scanning, together with the indexing templates and the indexes collected, will be stored on the structure of an electronic archive.

Searching (retrieving) indexed documents will use database search functions with client access.

#### 2.2.7. Training of the official - documentation for the use of the electronic archiving software and the resulting database

The training is done on the hardware and software platform used to implement the electronic archive. The training courses cover the topics required for initiation in the activities of Archive Administration (1 user) and Archive Use (x users). Handing over the electronically processed archival fund to the person in charge with the beneficiary's archive.

## 3. Conclusions

Electronic archiving has proven useful to all economic entities. Electronic archiving is used for daily electronic archiving, as a solution for storing and indexing documents, but also for electronically searching documents in the archive. In order to complete the archiving process, several steps are presented which are presented in this paper.

## 4. Reference

[1] F. Oprea, *Conținutul și evoluția conceptelor și ale practicilor de păstrare, conservare și restaurare a arhivelor*, Arhiva Românească, București, vol. 2, 1995, pp. 26-39.

[2] D. Dușmănescu, A. Pătrașcu, A. Tănăsescu, *A Data Model for Document's Electronic Archiving*, Rev. Inf. Econ., vol. 37, 2006, pp. 25-28.

[3] * * * - Legea Arhivelor Naționale nr. 16 / 1996 republicată în 2014.

[4] R. M. Pavelescu, *Arhivarea economică a documentelor*, Rev. In. Econ., vol. 10, 1999, pp. 72-75.

[5] R. M. Pavelescu, *Arhivarea economică componenta de bază a unui sistem de gestiune automată a documentelor*, Rev. In. Econ., vol. 19, 2001, pp. 40-43.

[6] L. Mera, *Îndreptar arhivistic*, Ed. Aronda, Cluj-Napoca, 2004.

**Bulletin of Scientific Information**
2019; 37
Published online JUNE 2019 (http://bsi.bioterra.ro)

# Bioinformatics in the study of molecular sequence alignments

**BETERINGHE Adrian[1], NICOLAE Marian[1]**

[1]Bioterra University of Bucharest, Romania, Faculty for Control and Expertise of Food, 013722, Bucharest, Garlei Street

**To cite this article:**
Beteringhe Adrian, Nicolae Marian, Colang George, Bioinformatics in the study of molecular sequence alignments. *Bulletin of Scientific Information*, No. 37, 2019, pp. 15-18.

**Abstract:** Sequential analysis, either local or global, is a basic tool of bioinformatics. Knowledge of biophysics, biochemistry, cell biology, probability theory and statistics but also computer science is required to define bioinformatics. At the base of the global alignment is the Needleman-Wunsch algorithm and in the case of local alignment the Smith-Waterman algorithm finds its usefulness.

**Keywords:** bioinformatics, global alignment, local alignment, Needleman-Wunsch algorithm, Smith-Waterman algorithm

## 1. Introduction

Although initially created as a discipline related to medical informatics [1], dedicated to the storage and processing of data from biology, subsequently, due to the accelerated growth of new information that contained more and more structural details, especially molecular sequences, new approaches were imposed in the organization. and systematization of data and efficient processing algorithms, but also the need to train specialists in bioinformatics.

To study bioinformatics, solid knowledge of biophysics, biochemistry, cell biology, probability theory and statistics, topology or series analysis is required. Also, systems theory, cybernetics and evolutionism theory contribute to an integrative vision, and databases and programming languages allow for easier management and processing of information in this field [2].

In bioinformatics, sequential alignment is a way of arranging DNA, RNA, or protein sequences in order to identify similar regions that may be consequences of functional, structural, or evolutionary relationships between sequences.

In order to be able to evaluate how similar two or more structures are, it was necessary to introduce methods and algorithms that allow a quantitative expression of the conclusions of the comparison [3].

As a rule, sequence comparisons are done with the purpose of identifying whether portions of these have been derived from common ancestors. The derivation is done through mutation and selection processes. The basic mutational processes are:

- substitutions: when one element is replaced by another;
- insertions: when one or more elements are inserted in a sequence;
- deletions: when one or more elements are deleted in a sequence.

The elements are represented by nucleotides in the case of DNA and RNA, respectively amino acids, in the case of proteins. Sequential alignment is achieved by writing sequences one by one (element by element), with the purpose of identifying similar portions.

During the alignment process there are three situations [4]:

- if in both sequences we have the same symbol on the same position, we will consider that the position has been preserved in evolution;

- if we have different symbols in the sequences, we will consider that they come from a common ancestor;
- if the sequences have different lengths, we will assume that there were possible insertions or deletions.

## 2. Results and conclusions

The problem that researchers have encountered over time is finding an algorithm for sequential analysis that allows the selection of the optimal alignment between two sequences, also taking into account a given scoring scheme. The algorithm created by Needleman and Wunsch in 1970 [5] is the best-known algorithm for the global alignment of two molecular sequences and can be applied to both proteins and nucleic acids.

To illustrate the application of the NW algorithm, we will consider the following two sequences:

X: A A T C

Y: A G C G

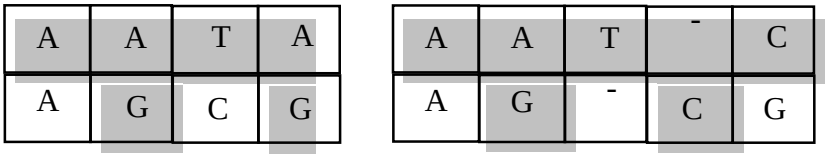
and the following score scheme (equation 1):

$$S = \begin{cases} +2, \text{math} \\ +1, \text{misfit} \\ -2, \text{gap} \end{cases} \quad (1)$$

**Table 1.** The matrix used in the NW algorithm as well as the specific "trace-back" algorithm

| | | **A** | **A** | **T** | **C** |
|---|---|---|---|---|---|
| | 0 | -2 | -4 | -6 | -8 |
| **A** | -2 | 1 | 0 | -2 | -4 |
| **G** | -4 | 0 | 2 | 1 | -1 |
| **C** | -6 | -2 | 1 | 3 | 3 |
| **G** | -8 | -4 | -1 | 2 | 4 |

Final Alignment:




**ST = 5** **ST = 0**

**Figure 1**. Optimal alignment for the considered example

In 1980, for Smith and Waterman local alignment, they introduced an algorithm [6] similar to the Needleman-Wunsch algorithm, but introduced the following changes:

• at the initialization instead of penalty for the gaps the edges will have only the value 0;

• in the calculation of a matrix element, the comparison will be made, as in the Needleman-Wunsch algorithm, the values obtained from the three neighboring boxes of potential origin (diagonal left-up, the box on the left, respectively the upper one), choosing the maximum value; but if all are negative, the allocation will be made with the value 0; In this way the alignment matrix will have only positive or 0 elements;

• at "trace-back" the starting point will not be started from the lower right corner, but from the highest value;

• the alignment can stop anywhere in the matrix, eventually reaching the upper or left edge if any of the aligned sequences show the area of interest right at the beginning.

To illustrate the application of the SW algorithm, we will consider the following two sequences:

X: A T T C

Y: A C C A

and the following score scheme (equation 2):

$$S = \begin{cases} +2, \text{math} \\ -1, \text{misfit} \\ -2, \text{gap} \end{cases} \quad (2)$$

**Table 2.** The matrix used in the SW algorithm as well as the specific "trace-back" algorithm

| | | **A** | **T** | **T** | **C** |
|---|---|---|---|---|---|
| | 0 | 0 | 0 | 0 | 0 |
| **A** | 0 | 2 | 0 | 0 | 0 |
| **C** | 0 | 0 | 1 | 0 | 0 |
| **C** | 0 | 0 | 0 | 0 | 2 |
| **A** | 0 | 2 | 0 | 0 | 0 |

Final Alignment:

| | | | |
|---|---|---|---|
| A | T | T | C |
| A | - | C | C |

**ST = 1**

**Figure 2**. Optimal alignment for the considered example

## 3. Conclusions

Computerized analysis of biological sequences has grown considerably over the last decade. The combination of applications for analyzing the alignments with the search tools in the current databases represents the evolutionary points that this field is targeting. The two algorithms prove to be efficient in all areas that require such sequential

analysis.

## 4. References

[1] G. I. Mihalaș, A. Tudor, S. Paralescu, *Bioinformatică*, Ed. Victor Babeș, Timișoara, 2011.

[2] A. M. Lesk, *Introduction to Bioinformatics*, (3rd edition), Oxford Univ Press, Oxford UK, 2008.

[3] W. R. Pearson, D. J. Lipman, *Improved tools for biological sequence comparison*, Proc. Natl. Acad. Sci. USA, 85, 1988, pp. 2444-2448.

[4] C. Lee, *Generating consensus sequences from partial order multiple sequence alignment graphs*, Bioinformatics 19, 2003, pp. 999–1008.

[5] S.B. Needleman, C.D. Wunsch, *A general method applicable to the search for similarities in the amino acid sequence of two proteins*, J. Mol. Biol. 48 (3), 1970, pp. 443–453.

[6] T. F. Smith, M.S. Waterman, *Identification of Common Molecular Subsequences*, J. Mol. Biol., 147, 1981, pp. 195-197.

**Bulletin of Scientific Information**
2019; 37
Published online JUNE 2019 (http://bsi.bioterra.ro)

# Multiple sequential alignment - the star model

**BETERINGHE Adrian[1], POPÎRLAN Alina Maria[1], Ciobanu ROXANA MARILENA[1]**

[1]Bioterra University of Bucharest, Romania, Faculty for Control and Expertise of Food, 013722, Bucharest, Garlei Street

**To cite this article:**
Beteringhe Adrian, Popârlan Alina Maria, Ciobanu Roxana Marilena, Multiple sequential alignment - the star model. *Bulletin of Scientific Information*, No. 37, 2019, pp. 19-22.

**Abstract:** Multiple sequential analysis is very important when one wants to identify similarities between several biological sequences. The two types of multiple alignment: progressive and iterative were the basis of the construction of star or tree models, models much faster in interpretations than adapting sequential alignment on pairs of several sequences.

**Keywords:** bioinformatics, global alignment, local alignment, Needleman-Wunsch algorithm, Smith-Waterman algorithm

## 1. Introduction

Multiple sequential alignment aims to identify similarities between multiple DNA sequences or amino acids (proteins). The similarity identified is all the more significant as it is true for several sequences: this suggests the presence of conserved regions within several evolutionary branches. The identification of multiple similarities is useful in designing experiments for testing and modifying the functions of specific proteins, in predicting the function and structure of proteins and in identifying new members in protein families [1].

As we noted in the previous subchapter, the blueberry gap method involves aligning two nucleotide / protein sequences. If we want to align three or more biological sequences, we call Multiple Sequence Alignment (MSA) [2].

There are two types of multiple alignment:

- ***Progressive,*** in which a beginning sequence is chosen which will be compared one at a time with the other;
- ***Iterative,*** which involves realigning the sequences over several iterations of the process.

Progressive MSA [3] is one of the fastest approximations, much faster than adapting sequential pairwise alignment of multiple sequences. A major disadvantage of this is that it is based on a good alignment of the first two sequences. Errors can spread throughout the entire multiple alignment.

In the case of multiple iterative alignment [3], the MSA is reiterated, starting with the alignment of the sequences in subgroups and ending with the subsequent alignment of the subgroups. Iterative MSA is an optimization method and can use genetic algorithms and hidden Markov models [4]. The disadvantage is that the process can be blocked locally and can become much slower.

Because for some multiple alignment methods (such as progressive alignment) the order in which the sequences are taken matters a great deal, we will next discuss sequence ordering patterns.

## 2. Results and conclusions

### 2.1. The "star" model

The star model [5] is one of the most used models for ordering biological sequences. This is done according to the following algorithm:

- there are given k sequences to be aligned: $s_1$, $s_2$,… $s_k$.

- a $s_c$ sequence is chosen as the center;

- for each sequence and $s_i \neq s_c$ is determined between and $s_i$ and $s_c$ an optimal alignment;
- the alignments of the pair meet;
- the result of multiple alignment is obtained by aggregating the pair alignments.

Note: In order to choose the center, each sequence is tried as a center and the best multiple alignment is obtained. Thus, a matrix of scores of dimensions (k + 2) × (k + 2) will be created, which will contain the k sequences on the first line and column. Since it is not necessary to compare a sequence with itself, the main diagonal will be filled with 0 [6], as in the example below:

| | $s_1$ | $s_2$ | -- | $s_k$ | Scoring |
|---|---|---|---|---|---|
| $s_1$ | 0 | | | | |
| $s_2$ | | 0 | | | |
| -- | | | 0 | | |
| $s_k$ | | | | 0 | |
| Scoring | | | | | -- |

**Figure 1**. The method of filling the matrix in the case of the star model

Next, the matrix will be filled with the Levenshtein distances calculated after the alignments have been made. In order to better understand the steps followed, we have the following example:

Example:

The sequences will be aligned: TGC, ATCC, AGTC, TGACGA, AATCGCA-, using the scoring scheme:

$$S = \begin{cases} +2, \text{math} \\ -1, \text{misfit} \\ -2, \text{gap} \end{cases} \quad (1)$$

The results presented in Figure 2 will be obtained:

| | $s_1$ | $s_2$ | $s_3$ | $s_4$ | $s_5$ | Scoring |
|---|---|---|---|---|---|---|
| $s_1$ | 0 | 6 | 6 | 1 | -9 | 4 |
| $s_2$ | 6 | 0 | 10 | -1 | -2 | 13 |
| $s_3$ | 6 | 10 | 0 | 2 | 1 | 19 |
| $s_4$ | 1 | -1 | 2 | 0 | 2 | 4 |
| $s_5$ | -9 | -2 | 1 | 2 | 0 | -8 |
| Scoring | 4 | 13 | 19 | 4 | -8 | -- |

**Figure 2**. The resulting matrix for the star model

It can be observed that the maximum score is 19, a value corresponding to the sequence $s_3$. This will be the center of $s_c$. The star model will look as shown below:

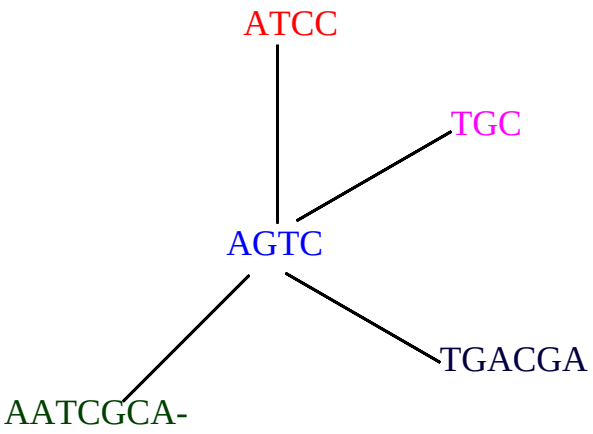



**Figure 3.** Sequence alignment according to the "star" model

Note: If, when aligning the center sequence with another, a gap appears inside it, when the final alignment is performed, a gap will be inserted in the same position and in the other sequences.

## 3. Conclusions

Multiple sequential analysis has proven useful for sequences that do not have the same length. The star model is important in multiple sequential analysis. Because sequence order is important in analysis, MSA models prove it most.

# 4. References

[1] C. Grasso, C. Lee, *Combining partial order alignment and progressive multiple sequence alignment increases alignment speed and scalability to very large alignment problems*, Bioinformatics, 20, 2004, pp. 1546–1556.

[2] L. Wang, Y. Jiang, *On the Complexity of Multiple Sequence Alignment*, J. Comp. Biol. 1, 2009, pp. 337-348.

[3] M. Maiolo, X. Zhang, M. Gill, M. Anisimova, *Progressive multiple sequence alignment with indel evolution*, *BMC Bioinformatics*, 19, 2018, pp. 331-338.

[4] S. Fine, Y. Singer, *The hierarchical hidden Markov model: analysis and applications*, Machine Learning, 32, 1998, pp. 41-62.

[5] Q. Zou, X. Shan, Y. Jiang, *A Novel Center Star Multiple Sequence Alignment Algorithm Based on Affine Gap Penalty and K-Band*, Phys. Procedia. 33, 2012, pp. 322–327.

[6] D. J. Lipman, S. F. Altschul, J. D. Kececioglu, *A tool for multiple sequence alignment*, Proc. Natl. Acad. Sci. U.S.A., 86, 1989, pp. 4412-4415.

**Bulletin of Scientific Information**
2019; 37
Published online JUNE 2019 (http://bsi.bioterra.ro)

# QSPR studies involving fullerene derivatives with importance for organic photovoltaic cells

**AMZOIU Emilia[1], AMZOIU Manuel[1], BETERINGHE Adrian[2]**

[1]University of Medicine and Pharmacy of Craiova, Romania, Faculty of Pharmacy, 200349, Craiova, Petru Rares Street
[2]Bioterra University of Bucharest, Romania, Faculty for Control and Expertise of Food, 013722, Bucharest, Garlei Street

**To cite this article:**
Amzoiu Emilia, Amzoiu Manuel, Beteringhe Adrian, QSPR studies involving fullerene derivatives with importance for organic photovoltaic cells. *Bulletin of Scientific Information*, No. 37, 2019, pp. 23-26.

**Abstract:** The QSPR (Quantitative Structure - Property) technique was used for the construction of a high-power model for predicting the parameters of a photovoltaic cell involving 4 n-type semiconductors from the fullerenes class.

**Keywords:** QSPR, fullerene derivatives, photovoltaic cell, filling factor, ArgusLab

## 1. Introduction

In 1996, the Nobel Prize in Chemistry was won by Harold W. Kroto, Robert F. Curl and Richard E. Smalley for the discovery in 1985 of a new allotropic form of carbon in which atoms are arranged in closed forms. The new structure is shaped like a truncated icosahedron and was named Buckminsterfullerene (C60) by the name of architect Buckminster Fuller who in 1960 designed several geodesic domes [1]. After the discovery of C60, many experimental studies have focused on the synthesis, isolation, characterization [2-3] as well as the study of geometry [4] or electronic structure [5] for C60.

C60 consists of 60 carbon atoms arranged in 12 pentagons and 20 hexagons. The structure of Buckminsterfullerene resembles a soccer ball with the black areas being the pentagons and the white ones being the hexagons. Among the spectacular properties that the C60 possesses are the following:

i) the ability to accept up to 6 electrons reversibly (electronegative molecule) [6];

ii) superconductor in species $M_3C60$ (M = alkali metal) [7];

iii) material with nonlinear optical properties [8].

The photovoltaic effect in organic solar cells has been studied since 1950 when voltages around 1V were obtained on sandwich structures from organic thin layers. The modification of the chemical structures for the organic compounds has led to the increase of the efficiency of these photovoltaic devices. The first organic photovoltaic cells were made using M1-organic semiconductor (SO) - M2 sandwich structures, in which M1 is a thin semitransparent thin layer of metal with a small extraction work (eg Al) that forms a contact with SO. blocking and M2 is a semi-transparent thin layer of a metal with a high extraction work (eg Au, Ag, Cu) that forms an ohmic contact with SO. It can be used as an organic semiconductor, C60, which acts as a n-type semiconductor. Instead of metals,

electron-donor polymers (polythiophene) can be used which are p-type semiconductors. By forming them into thin layers a heterojunction is formed.

In photovoltaic cells, both fullerenes as such and derivatized are used. The most commonly used is Buckminsterfulerena C60 but studies have shown that fullerene C70 is more efficient in energy conversion by 25% [10]. Figure 1 shows the structures of the C60 and C70 fullerenes respectively.

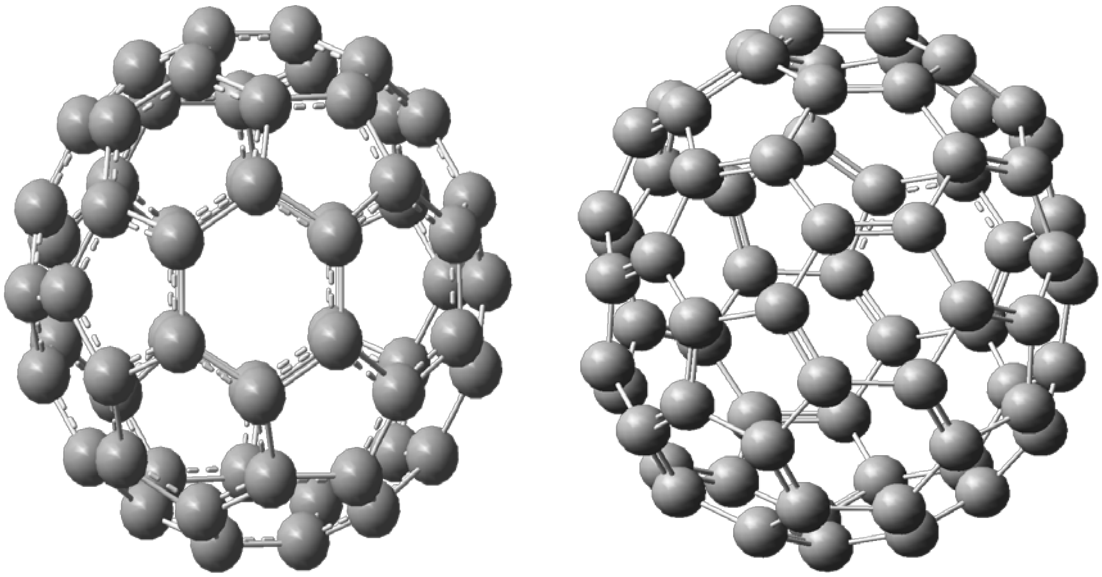


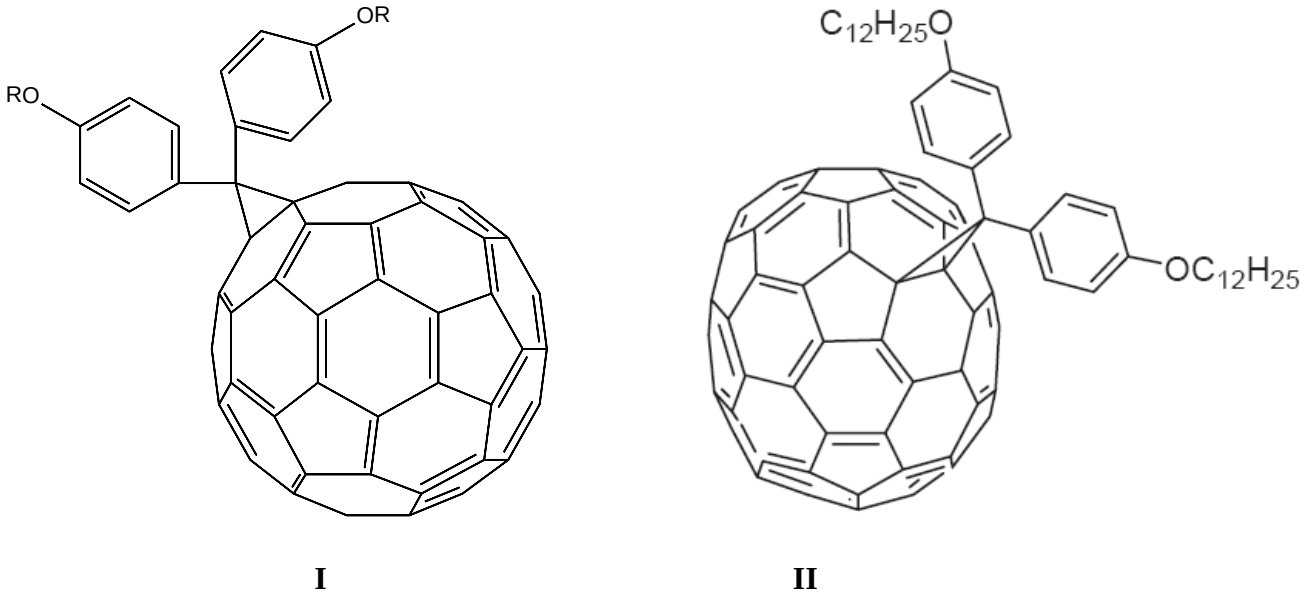

**Figure 1**. Structures of the C60 and C70 fullerenes [11]

In this study, 4 fullerene derivatives (1-4) were used, of which 3 of class C60 (I) and one derivative of class C70 (II) whose values of the filling factor (FF) [12] were involved in the construction. a QSPR model having as molecular descriptors those generated by the ArgusLab program [13].

.

## 2. Results and conclusions

Figure 2 shows the structures of the 4 fullerene derivatives (1-4) [12].




**I** **II**

**a:** $R^{-}$ —— $CH(CH_3)_2$

**b:** $R^{=}$ —— $C_6H_{13}$

**c:** $R^{=}$ —— $C_{12}H_{25}$

**Figure 2**. Structure of derivatives involved in the QSPR study

The optimal equation resulting from the multilinear regression involving 3 descriptors for the FF filling factor is:

$$FF = 9.754 \cdot 10^{-4}(\pm 1.740 \cdot 10^{-5})NA - 1.167(\pm 0.223)RNDB - 0.264 \quad (1)$$

$R^2 = 0.901$ SD = 0.098 F = 22.45 $R^2(CV) = 0.805$

where: NA represents the number of atoms and RNDB represents the relative number of double bonds. The 2 descriptors were calculated using the ArgusLab program [13].

## 3. Conclusions

The statistical parameters as well as the cross-validation of the model (R2 (CV)) (equation 1) make this model to be used for the prediction of the filling factor and for other fullerene derivatives to be designed and synthesized.

.

## 4. References

[1] H. W. Kroto, J. R. Heath, S. X. O'Brien, R. F. Curl, R. F. Curls, *$C_{60}$: Buckminsterfullerene*, Nature, 318, 1985, pp. 162-163.

[2] G. Flamant, D. Luxemburg, J. F. Robert, D. Laplaze, *Optimizing fullerene synthesis in a 50 kW solar reactor*, Solar Energy, 77, 2004, pp. 73-80.

[3] H. Ohta, Y. Saito, N. Nagae, J. J. Pesek, M. T. Matyska, K. Jinno, *Fullerenes separation with monomeric type C30 stationary phase in high-performance liquid chromatography*, J. Chrom. A, 883, 2000, pp. 55-66.

[4] B. L. Zhang, C. Z. Wang, K. M. Ho, C. H. Xu, C. T. Chan, *The geometry of small fullerene cages: $C_{20}$ to $C_{70}$*, J. Chem. Phys., 97, 1992, pp. 5007-5011.

[5] F. Jensen, H. Koch, *C24: Ring or fullerene?*, J. Chem. Phys., 108, 1998, pp. 3213-3217.

[6] L. Echegoyen, L. E. Echegoyen, *Electrochemistry of Fullerenes and Their Derivatives***,** Acc. Chem. Res., 31, 1998, pp. 593-601.

[7] R. C. Haddon, A. F. Hebard, M. J. Rosseinski, D. W. Murphy, S. J. Duclos, K. B. Lyons, B. Miller, J. M. Zahurak, R. Tycko, G. Dabbagh, F. A. Thiel, *Conducting films of $C_{60}$ and $C_{70}$ by alkali-metal doping***,** Nature (London), 350, 1991, pp. 320-322.

[8] L. W. Tutt, A. Krost, *Optical limiting performance of $C_{60}$ and $C_{70}$ solutions*, Nature (London), 356, 1992, pp. 225-226.

[9] S. Antohe, *Materiale şi dispozitive electronice organice*, Ed. Universităţii din Bucureşti, 1996.

[10] B. C. Yadav, R. Kumar, *Structure, properties and applications of fullerenes*, Internat. J. Nanotech. Appl., 2008, 2, pp. 15-24.

[11] J. B. Howard, J. T. McKinnon, Y. Makarovsky, A. L. Lafleur, M. E. Johnson, *Fullerenes $C_{60}$ and $C_{70}$ in flames*, Nature, 352, 1991, pp. 139-141.

[12] J. L. Segura, F. Giacalone, R. Gomez, N. Martin, D. M. Guldi, C. Luo, A. Swartz, I. Riedel, D. Chirvase, J. Parisi, V. Dyakonov, N. S. Sariciftci, F. Padinger, *Design, synthesis and photovoltaic properties of [60] fullerene based molecular materials*, Materials Sci. Eng. C, 2005, 25, pp. 835-842.

[13] D. Feller, M. A. Thompson, R. A. Kendall, A Theoretical Case Study of Substituent Effects and Microsolvation on the Binding Specificity of Crown Ethers, J. Phys. Chem. A, 101, 1997, pp. 7292-7298.

**Bulletin of Scientific Information**
2019; 37
Published online JUNE 2019 (http://bsi.bioterra.ro)

# The utility of molecular docking technique in the study of Boswellia resin (Frankincense)

**BETERINGHE Adrian[1], AMZOIU Emilia[2], AMZOIU Manuel[2], POPESCU Sofia[3]**

[1]Bioterra University of Bucharest, Romania, Faculty for Control and Expertise of Food, 013722, Bucharest, Garlei Street
[2]University of Medicine and Pharmacy of Craiova, Romania, Faculty of Pharmacy, 200349, Craiova, Petru Rares Street
[3]Banat's University of Agriculture Science and Veterinary Medicine, 300645, Timişoara

**To cite this article:**
Beteringhe Adrian, Amzoiu Emilia, Amzoiu Manuel, Popescu Sofia, The utility of molecular docking technique in the study of Boswellia resin (Frankincense). *Bulletin of Scientific Information*, No. 37, 2019, pp. 27-30.

**Abstract:** Boswellia resin known in history but also in medical studies as incense has proved useful not only in religious ceremonies but also in the medical field through the chemical compounds contained in its composition. Boswellic acids prove potential inhibitors of COX-1, a fact proven both theoretically by Molecular Docking and experimental technique.

**Keywords:** Boswellia resin, Frankincense, Olibanum, Molecular Docking Technique, COX-1, COX-2, Boswellic acids

## 1. Introduction

The incense is the resin of some species of shrubs of the genus Boswellia from India, Somalia and Southern Arabia (Oman) which - through warming, emit a pleasant smell of balsam. To obtain incense, the bark of the trees is manually cropped and the sap from the inside, a white liquid with a creamy texture, is collected only after 2 weeks and then allowed to dry. Boswellia serrata trees produce quality resin 3 years in a row, after which it is important that the sap harvesting procedure be interrupted for several years in order for them to regenerate [1, 2].

The incense is mentioned in one of the oldest known medical archives, the Eber papyrus (dating from the 16th century BC), but also a series of prescriptions and medical prescriptions from Ancient Egypt [3].

Boswellia (Frankincense, Olibanum) resin has been used as a size in religious ceremonies since the beginning of history [4]. From a medical point of view the size has been used due to the anti-inflammatory activity or the anti-microbial activity of the chemical components of the resin and especially of the α, β boswellic acids [5, 6].

The structure of β-boswellic acid is shown in Figure 1 [1, 6]:

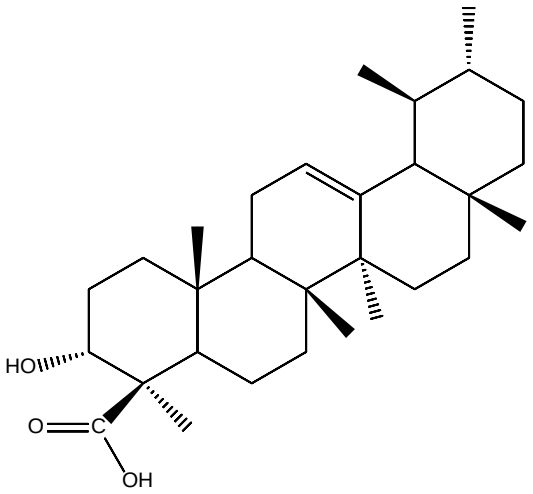



**Figure 1.** Structure of β-boswellic acid [1, 6]

Later in the literature it has been shown that incensol acetate (another important component of Bowellia or Frankincense resin) exerts an important neo-protective effect after brain injury in mice. It also has an antidepressant and anxiolytic effect in mice [7].

The application of computational methods in studying the formation of intermolecular complexes is the subject of intense research in recent years. One of these methods called "Molecular Docking" can be used to predict the favorable orientations of one molecule (ligand) to a second molecule (receptor) such that by binding the two molecules to form a stable complex [8].

Knowing the preferred orientations can be used to predict the binding affinity between two molecules using the scoring functions. This method is also frequently used in predicting the affinity of binding of small molecules (drugs) to target proteins or the biological activities of these molecules (drugs) [9].

In this paper we will focus on boswellic acids and their effects using the Molecular docking technique.

The Molecular Docking technique has been found useful in testing boswellic acids as potential inhibitors of cyclooxygenase-1 (COX-1) (3N8V [10, 11], Figure 2), an enzyme that catalyzes the conversion of arachidonic acid into Prostaglandin H2, via - a short-lived Prostaglandin G2 intermediate [10].

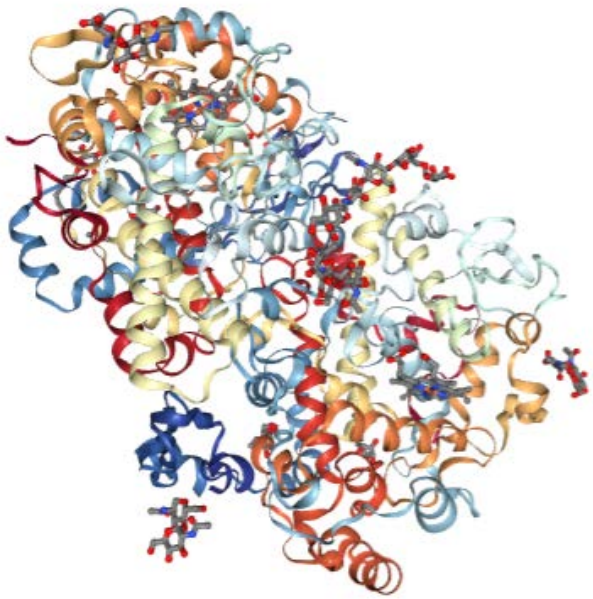


**Figure 2**. Cyclooxygenase-1 (COX-1) structure (code 3N8V-Protein Data Bank [10])

## 2. Results and conclusions

Both the COX-1 structure (code 3N8V-Protein Data Bank [10]) as well as the boswellic acid structures were prepared for the “molecular docking” process using the Hex 8.0 program [12], being modeled using 3D parametric coding functions. as well as surface shape, electrostatic charge as well as potential distribution. These parametric functions are based on spherical or polar orthogonal basis functions.

Because by translating and rotating the ligand, the "docking" process is more complete in Table 1, the binding energy values (affinities) (kJ / mol) of boswellic acid-COX-1 complexes will be presented.

**Table 1.** Binding energy values (affinities) (kJ / mol) of boswellic acids-COX-1 complexes

| No | α-boswellic acid | β-boswellic acid |
|---|---|---|
| COX-1 | -327.78KJ/mol | -329.60 KJ/mol |

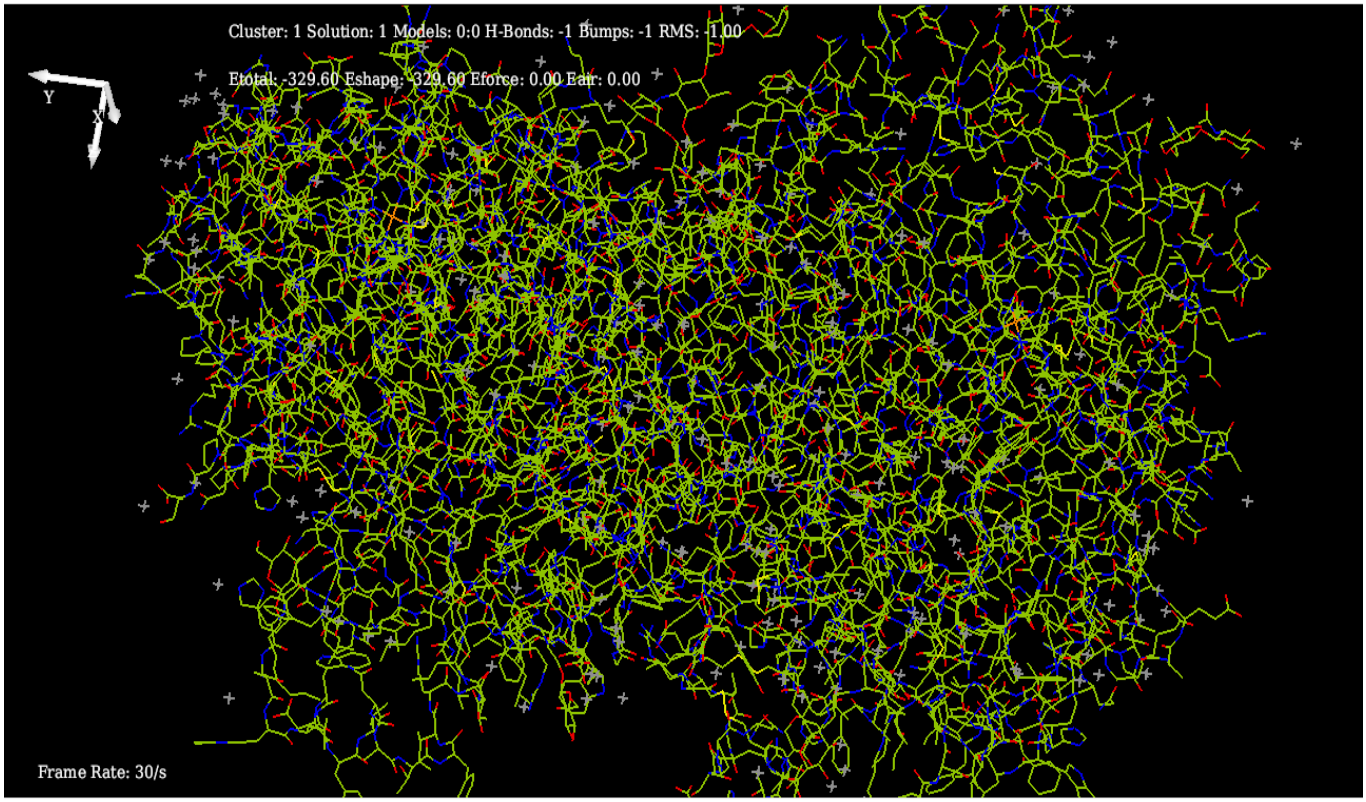



**Figure 3**. Structure of the β-boswellic acid complex - cyclooxygenase-1 (COX-1)

Studies have shown that boswellic acids have lower inhibitory activity than COX-2 compared to COX-1 [12].

## 3. Conclusions

The resin of Boswellia species ('Frankincense', 'Olibanum') through chemical compounds contained in its structure has proved useful not only in religious ceremonies but also for medical purposes. Molecular docking technique is useful in the study of boswellic acids as COX-1 cyclooxygenase inhibitors. From the data presented in Table 1 it can be concluded that β-boswellic acid is more efficient in inhibiting COX-1 but not as efficient in relation to COX-2.

.

## 4. References

[1] A. Moussaieff, R. Mechoulam, *Boswellia resin: from religious ceremonies to medical uses: a review of in-vitro, in-vivo and clinical trials*, J. Pharm. Pharmacol., 61, 2009, pp. 1-13.

[2] C. C. delasalle, M. Mehiri, C. Cagliero, P. Rubiolo, C. Bicchi, U. J. Meierhenrich, N. Baldovini, *The (+)-cis- and (+)-trans-Olibanic Acids: Key Odorants of Frankincense*, 55, 2016, pp. 13719-13723.

[3] T. Bardinet, *Les papyrus medicaux de l'Egypte pharaonique*, ed. Fayard, Paris, 1995.

[4] S. Marshall, *Frankincense: festive pharmacognosy*, Pharm. J., 271, 2003, 862-864.

[5] S. Hamm, J. Bleton, J. Connan, A. Tchapla, *A chemical investigation by headspace SPME and GC-MS of volatile and semi-volatile terpenes in various olibanum samples*, Phytochem., 66, 2005, pp. 1499-1514.

[6] H. P. Ammon, *Boswellic acids in chronic inflammatory diseases*, Planta Med., 72, 2006, pp. 1100-1116.

[7] A. Moussaieff, J. Yu, H. Zhu, S. Gattoni-Celli, E. Shohami, M. S. Kindy, *Protective effects of incensole acetate on cerebral ischemic injury*, Brain Res., 1443, 2012, pp. 89-97.

[8] T. Lengauer, M. Rarey, *Computational methods for biomolecular docking*, Curr. Opin. Struct. Biol., 6(3), 1996, pp. 402-406.

[9] D. B. Kitchen, H. Decornez, J. R. Furr, J. Bajorath, *Docking and scoring in virtual screening for drug discovery: methods and applications*, Nature reviews. Drug discovery, 3(11), 2004, pp. 935-949.

[10] https://www.rcsb.org/structure/3N8V

[11] R. S. Sidhu, J. Y. Lee, C. Yuan, W. L. Smith, *Comparison of Cyclooxygenase-1 Crystal Structures: Cross-Talk between Monomers Comprising Cyclooxygenase-1 Homodimers*, Biochem., 49, 2010, pp. 7069-7079.

[12] D. Ritchie, Hex Protein Docking, http://www.csd.abdn.ac.uk/hex/

**INTERNATIONAL PARTNERS**

**REPUBLIC OF MOLDOVA**

State University "Dimitrie Cantemir" of the Academy of Sciences of Moldova

**EGIPT**

Khadiga Mohamed Gaafar - Higher Institute for Specific Studies
May Mohamed El Batran - Higher Institute for Specific Studies
Mohamed Mahmout El Batran - Higher Institute for Specific Studies

**ITALIA**

Balestri Gino - D.A.G.A. Sezione Scienze Zootecniche, Universita di Pisa
Doni S. - Istituto per lo Studio degli Ecosistemi (ISE), Pisa
Fantoni Elena - Comune di San Giuliano Terme Pisa
Fontanelli Marco - University of Pisa
Fratini Roberto - Department of Florence University
Gajo Paolo - Florence University
Ginanni M. - University of Pisa
Lorenzini G. - Dipartimento di Scienze Zootecniche, Universita di Firenze
Macci C. - Istituto per lo Studio degli Ecosistemi (ISE), Pisa
Mani Danilo - D.A.G.A. Sezione Scienze Zootecniche, Universita di Pisa
Martini Andreea - Dipartimento di Scienze Zootecniche, Universita di Firenze
Masciandaro Graziana - ISE-CNR Pisa
Migliorini P. - Universita degli Studi di Scienze Gastronomiche
Peruzzi A. - Centro Interdipartimentale di Ricerce Agro-Ambientali "Enrico Avanzi", University of Pisa
Peruzzi E. - Istituto per lo Studio degli Ecosistemi (ISE), Pisa
Pistoia Alessandro - D.A.G.A. Sezione Scienze Zootecniche, Universita di Pisa
Poli Piera - D.A.G.A. Sezione Scienze Zootecniche, Universita di Pisa
Raffaelli M. - University of Pisa
Siboni Eugenio - Societa Produttori Sementi SpA, International Marketing Manager, Argelato (BO)