



Revista Cubana de Química

ISSN: 0258-5995

revcubanaquimica@cnt.uo.edu.cu

Universidad de Oriente

Cuba

Agüero Chapin, Guillermin; González Díaz, Humberto; Molina, Reinaldo; Uriarte, Eugenio
"2D RNA-QSAR: Assigning Acc Oxidase Family Membership with Stochastic Molecular
Descriptors. Isolation and Prediction of a Sequence from Psidium Guajava L"

Revista Cubana de Química, vol. XVII, núm. 2, 2005, p. 227

Universidad de Oriente

Santiago de Cuba, Cuba

Available in: <http://www.redalyc.org/articulo.oa?id=443543686051>

- How to cite
- Complete issue
- More information about this article
- Journal's homepage in redalyc.org

redalyc.org

Scientific Information System

Network of Scientific Journals from Latin America, the Caribbean, Spain and Portugal

Non-profit academic project, developed under the open access initiative

“2D RNA-QSAR: Assigning Acc Oxidase Family Membership with Stochastic Molecular Descriptors. Isolation and Prediction of a Sequence from *Psidium Guajava* L”

Guillermin Agüero Chapin¹ Humberto González Díaz², Reinaldo Molina², and Eugenio Uriarte³

¹Process Analysis Center (CAP), Faculty of Chemistry and Pharmacy, Central University of ‘Las Villas’, Carretera a Camajuaní Km 5/2, Santa Clara, Cuba, CP: 54830 email: chapin@gf.uclv.edu.cu.

²Chemical Bioactive Center (CBQ), Central University of ‘Las Villas’ 54830, Cuba

³Department of Organic Chemistry, Faculty of Pharmacy, University of Santiago de Compostela 15782, Spain

Abstract

Quantitative structure–activity relationship (QSAR) techniques for small molecules could be applied to nucleic acids. Unfortunately, almost all molecular descriptors are more successful at encoding branching information than sequences and/or cannot be back-projected. A solution for scaling the QSAR problem up to RNA may be to transform sequences into secondary structures first. Our group has used Markovian negentropies as molecular descriptors for drug design with preliminary results in 20 bioinformatics [Bioinformatics 2003, 19, 2079]. However, RNA-QSAR studies on RNA molecules have not been described to date. Novel Markovian negentropies have been introduced here as molecular descriptors for 2D-RNA structures. An RNA-QSAR study of the ACC proteins from different plants has been carried out. The QSAR recognizes 19/20 sequences (95.0%) within the ACC family and 12/17 (70.6%) of the control group sequences. The model has a high Matthews regression coefficient ($C = 0.68$). Overall cross-validation average accuracies were 14 out of 15 for ACC sequences (93.3%) and 10 out of 13 for control sequences (76.9%). Finally, ACC oxidase family membership was assigned to a new sequence isolated for the first time in this work from *Psidium guajava* L. A backprojection map for this sequence identifies the left stem (40%) and the main stem (45%) as highly important substructures. Results of an nBLAST experiment are consistent with this finding and indicate a high conservation score (>70) for left stem and main stem; whereas major loop, right

stem, cap and major loop right half were hardly conserved.

1. Vázquez-Padrón, R. I.; de la Riva, G.; Agüero, G.; Silva, Y.; Pham, S. M.; Soberon, M.; Bravo, A.; Aït-touche, A. FEBS Lett. 2004, 578, 30.
2. González-Díaz, H.; Uriarte, E.; Ramos de Armas, R. Bioorg. Med. Chem. 2004, 13, 323.

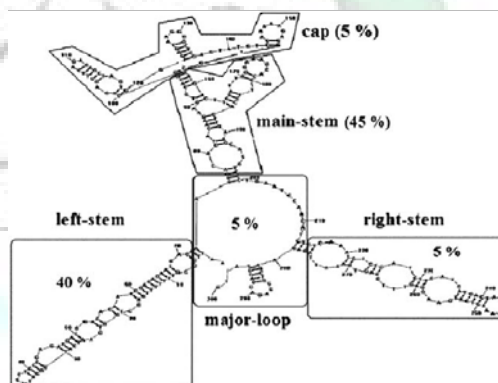


Fig 1. Backprojection map for guava ACC oxidase RNA secondary structure