Available online at www.sciencedirect.com
SCIENCE DIRECT®
Il Farmaco 58 (2003) 621–623
IL FARMACO
www.elsevier.com/locate/farmac

# QSAR and the role of luck in research

It is a great pleasure to honor Professor Pratesi and speak to you about QSAR. Professor Pratesi was one of the first in the world to understand the importance of QSAR. Three years after our first publication, he invited me to give my first formal report at a meeting in Sao Paolo, Brazil in 1965. I appreciated this opportunity in spite of being criticized by a Merck researcher who suggested that what we were trying to do was ridiculous! Professor Pratesi, I want to add, was instrumental in my being awarded the Medal of the Italian Society of Pharmaceutical Sciences in 1976.

Scientists sometimes have the feeling that they really are in charge and plan their future step-by-step. Not always the case. Looking back on the evolution of QSAR, I see it was largely a matter of luck. The first lucky event occurred in my first year in a small junior college in Lincoln, IL for there, I met Fred Uhle, an unusual student. He had worked his way through an Organic Chemistry textbook in high school. He interested me in the subject and we soon set up a lab in the basement of my house and started making various dyes. When he graduated from the University of Illinois, a year ahead of me, he went to Columbia University in New York. I had two options after graduating: an assistantship at the University of California or New York University. Naturally, I chose New York in order to work with Fred. This determined the rest of my life (it was here, I might mention, that I met my wife—more about her later).

At NYU, one of my best friends was Warren Garrison. Upon graduating, he took a research position at the University of Chicago. The next year, I went back to my alma mater, the University of Illinois to work on a government project on antimalarials. Soon, with World War II was winding down, the government discontinued the project and I was left without a job. I contacted Garrison, who said he was sure I could get a position at the University of Chicago, but he could not reveal the nature of the research! I had to accept the position on faith. He assured me that the work was very interesting.

On day 1, I learned an atomic bomb was being developed and, for example, if one were dropped on London, the city would be completely destroyed. This was a terrible shock as I had never even contemplated the idea of such a bomb.

I was lucky to meet Nelson Smith on this project. Soon, we were both transferred to Richland, WA, where the production of plutonium was about to start. My job was to develop a method for analysis that housewives could use. After about a year, no further research was needed and we had to look for new jobs. Smith took a position at Pomona College, his alma mater. I had never heard of Pomona College and told Smith he was foolish to go to a tiny school that had no graduate students and hence, no potential for research. With the war still going on and few teaching positions, I accepted an offer from the duPont Company.

I was the one who made a terrible mistake, ending up working under a man called by his postdocs "Wild Bill Gresham". He required us to write daily reports on our work, then summarize them in weekly reports and finally to write monthly reports. Of course, when a project was finished, another report was due. Every morning he came around to find out our plans for the day and often checked up on the progress in the afternoon. In the 3 months I could stomach all this, I finished three projects—AND decided to get out of that pressure cooker and find another job, any job!

As luck would have it, Nelson Smith informed me that Pomona College needed an Organic Chemistry teacher. So I crossed the country to go to the school I had thought Nelson was crazy to choose, hoping from there to look for a new position. The College was located in a town of 4000 people with fewer than 1000 students; however, it was so peaceful and generally attractive that I never thought of leaving. Forget about great research at a big University!

On arriving at Pomona, I met Robert Muir, a plant physiologist who certainly changed my life. For lack of space in the biology building, he was working in the chemistry building. We soon discovered that we had a common interest in plant growth regulators. I had been working in graduate school on indole acetic acid, an important growth regulator. Another fortuitous incident! At that time, 2,4-dichlorophenoxyacetic acid was an important weed killer, but at low concentrations it promoted plant growth. Soon I was busy making all kinds of phenoxyacetic acids that Muir tested on sections of oat shoots. It slowly appeared to us that one needed an open ortho position and an electron-

0014-827X/03/$ - see front matter 
doi:10.1016/S0014-827X(03)00175-7

attracting substituent in the meta position; substituents in the para position blocked activity.

About 1950 our department chairman asked me to give an advanced course in synthetic organic chemistry, my specialty. However, I had just accidentally discovered in our library, a book by Remick that really changed my life. The book was entitled, ''Electronic Interpretations of Organic Chemistry''. For some strange reason, I became intensely interested in mechanistic physical organic chemistry and made this the subject of my advanced course. I knew so little about the subject that I did not feel capable of lecturing on it. So for the first 2 years, I assigned different topics to the students on which to give oral reports. This would never have happened at a University where the course would have been assigned to a specialist! By 1961, I had a rough idea about the Hammett equation, but no idea about deriving equations with more than one term.

At this point in time, a College trustee who was starting to manufacture a miserably small computer, gave one to our department. Of course I had no idea how to use it, but as luck would have it, Donald McIntyre in our Geology Department had purchased one a year earlier and learned how to use it and he showed me how. One needed three specially wired boards. Enter the data, then invert the matrix and finally print out the results. The process required the better part of 1 h.

By this time I had learned from reading the literature that since the turn of the century, researchers had been trying to rationalize toxicity of chemicals to tadpoles and had found that the fatty character of chemicals defined by their partitioning between olive oil and water, was related to their toxicity to tadpoles. This substance was too viscous and hard to work with, so I decided to try octanol. That soon became the gold standard. We had discovered this shortly before Toshio Fujita arrived from Japan.

Getting Fujita to Pomona was an amazing piece of good fortune. In the mid-1950s a Japanese physicist Fukui, published an article saying that our simple idea about an electron-attracting substituent in the meta position was right. Later, when I got my first NIH grant with funds for a post-doc, I realized that I could not attract such a person from an American university to an undergraduate institution like Pomona College. Remembering Fukui's comments, I wrote him asking if he knew anyone in Japan who might be interested in the position. He knew Toshio Fujita, who was already working on plant growth regulators. After 2 years with us, Toshio was a confirmed researcher on QSAR. He eventually published more papers on the subject than any other laboratory except ours. Soon he was busy measuring partition coefficients in our laboratory. This led to the famous bumper sticker that a chemist spotted on a freeway near New York City: ''Hansch Walks on Octanol''.

The really difficult problem now began to become apparent. We could see qualitatively that an electronic factor must also be involved. My initial knowledge about the Hammett parameters provided this clue, but it was clear from simple inspection that the hydrophobic effect related to log P was not linear. In fact, this had stymied the early efforts of others to interpret the results with tadpoles and other simple organisms. I finally decided that we needed an equation parabolic in log P and linear in an electronic term. However we had no idea of how to formulate such an expression and worse yet, I had no knowledge of statistics. The simple Clary computer that was given to us could only handle a three-term expression. The next problem was statistics and at that time chemists, even physical organic chemists, did not understand how to use statistics. In fact, most of their studies could be dealt with via a simple plot of the data. A friend in the math department helped us deal with this problem.

Hammett parameters are a most important subject understood even today by few people studying the interaction of chemicals with biological systems.

Without electronic parameters serious work on QSAR is not possible. Quantum chemical parameters have been used by us, as well as others, but they are very time consuming and are often difficult to understand mechanistically. Our present database of over 10 000 bio QSAR contains 2400 equations based on Hammett terms that can be compared with 8800 equations that we have from physical organic chemistry based on these terms. In addition, we have 2500 bio QSAR that require terms for polarizability. Thus, almost half of the bio QSAR require terms for the electronic interactions. Still, most researchers consider Hammett parameters to be passé and essentially no one outside our laboratory considers polarizability. At a recent meeting about software development for biological studies, the director asked the group of about 80 people how many of them had heard of the Hammett equation. Four hands were raised. He then asked one person what it meant to him. He had no idea—he had only heard of it. Forget polarizability!

Another amazing piece of good luck occurred in 1968 when Al Leo joined our group. After receiving his Ph.D. at the University of Chicago, he was offered a position at the Shell Company. As luck would have it, upon administering the medical exam they discovered he had a metal plate in his head—the result of having been shot by a sniper in World War II (if the sniper had aimed a little bit to the left, Leo would have been killed). Shell would not take a chance on such a person, so Leo decided to manage the family company. After a few years, he became bored and moved back to the family house a few miles from Claremont. For lack of a job, he

decided to work with us to really gain a deeper understanding of log P. For a number of years after this we had a person working full time measuring log P for various types of chemicals to characterize their hydrophobicity. Gradually, Leo and I began to try to understand the molecular components behind log P. Slowly he developed a method of calculating log P that today is used all over the world.

Another event out of the blue occurred when David Weininger decided to join our group. We had no way to pay him, but the Environmental Protection Agency where he had been working, paid his salary for a while until we could support him from the sale of tables of log P values that we had been measuring and collecting from the literature. During his stay at Pomona, he developed the SMILES language for entering 2-dimensional structures of organic compounds in the computer system in one dimension. We could not have made any real progress without this language. There are other means for doing this now, but we modified his programs with places for the automatic loading of substituents. It would never have been possible to derive QSAR by the thousands without this feature. Very few people realized the great importance of SMILES in QSAR.

Up to this point in the early 1980s we had slowly begun to formulate a few hundred QSAR. Much of the early work had been done with an IBM computer that required cards for each data point that we processed by the computer. We had a large filing cabinet full of cards.

An astounding event occurred in 1987 when the R.J. Reynolds Tobacco Company decided, without any solicitation by us, to support our research. In fact, during the period from 1987 to 1994 they provided us with $650 000. It was one of the most crucial events in the development of QSAR.

At this time we needed someone to construct a really sophisticated computer program to derive QSAR via autoloading of parameters and then storing them in such a way as to make all the data available in myriad different ways. As chance would have it, at this moment a former Pomona student, David Hoekman, visited the College. He was designed in heaven for our project as at Pomona, he had spent 3 years studying mathematics and physics and then 2 years studying biology. We got in touch with him via Donald McIntyre who managed our tiny computer center as well as being Chairman of the Geology Department. Hoekman decided against working for a large company and has worked with us for 16 years. The evolution of our C-QSAR program to its present state needed someone with the dual biology–math background. He has also been very helpful in the development of the Clog P program that BioByte Corporation has supported for many years.

When I began getting grants from NIH, I had the problem of getting postdoctoral help. Good Ph.D.s from a first class university would not think of coming to a tiny undergraduate—only institution. The first stroke of good luck was getting Hua Gao who was finishing his Ph.D. at the University of Southern California with Professor Lien, one of the first people to work with me in 1968. He developed and entered into our system 4153 QSAR. When he decided to look for a regular job, I had more extremely good luck. Professor S.P. Gupta at the University of Pilani in India, who was interested in QSAR and had actually published a number of papers in this field, got in contact with me. From his laboratory over the years I got three outstanding people who developed and entered many equations: Dr. Rajni Garg (2046); Dr. Alka Kurup (2157) and Dr Suresh Mekapati (2481). All of this work was done after 1996.

A very important contributor to our work was Cynthia Selassie, who was born in Mombasa, Kenya, but on her own won a scholarship to a small College in Los Angeles. She worked with me for years before becoming a faculty member and eventually, Department Chair. And, to reiterate, one of the most helpful people to promote the use of QSAR was Professor Toshio Fujita.

Now with such a large database, we constantly are making novel discoveries. For instance, we have found a completely new way of finding allosteric reactions purely by chance. Up to this point in time, this was done by making studies on a single compound reacting with a given receptor at various concentrations and noting the inflection point. What we found was in many data sets, we had QSAR with a negative log P or CMR term and a positive such term squared. That is an inverted parabola. At first, activity decreased with increasing size and log P or CMR, then at an inflection point, turned around and increased. We now have dozens of such QSAR. Another interesting discovery is the simple addition of the number of valence electrons on the elements in a molecule to yield a parameter that is surprisingly effective in accounting for the polarizability factor in chemicals reacting with nerve synapses.

Final credit where credit is due: To Professor Pratesi for his early-on and continued support of my work, and to Gloria, my understanding wife, for managing so many matters in our day-to-day lives that I had the time necessary to develop our huge system.

Corwin Hansch
*Department of Chemistry, Pomona College, 645 North College Avenue, Claremont, CA 91711-6338, USA*
E-mail address: atessier@pomona.edu