ELSEVIER

Current status and future direction of the molecular modeling industry

Allen B. Richon

Molecular Solutions, Inc., 1116 Miller Mountain Road, Saluda, NC 28773, USA

During the period from 1970 through 1989, over 25 companies with products dedicated to computational chemistry were founded. By 1990, market forces within their organizations and the customer base they served began to change and these companies were forced to adapt in ways that were not anticipated.

Relative to other information-based industry segments, the computational chemistry 'industry' evolved over a very short period of time. During the 1960s, a limited number of scientists began exploring the use of mathematical techniques to derive relationships between observed activity and chemical structure. Between 1970 and 1990, several pharmaceutical and agrochemical organizations added computational chemists to their research departments and the growth of the commercial molecular modeling industry began. By 2007, computational chemistry was fully integrated into almost every aspect of the chemical research enterprise. Projects ranging from research into the structures of G-protein-coupled receptors [1] to the prediction of structural features necessary to optimize the activity of drug candidates [2] were routinely applied to drug discovery and development projects. The pharmacophore [3], a concept pioneered by generations of medicinal chemists was used in a variety of research programs that spanned disciplines from QSAR [4,5] to molecular similarity/combinatorial chemistry [6]. The full range of research projects that have been addressed by computational chemistry, while beyond the scope of this article, has been reflected by the submissions published in the 27 volumes of the *Annual Reviews in Computational Chemistry* [7].

With a history of such rapid growth, it was not surprising that the discipline attracted the interest of the financial community, which felt that the potential leverage of the technology would mirror that seen in the computer-aided design/manufacturing (CAD/CAM) marketplace. Meanwhile, changes within the user community, added to the problem of market saturation, forced many of the commercial modeling companies to redefine their initial strategies. At the same time, several computational groups within the chemical industry began to redefine their role from that of software developer to an emphasis on therapeutic project support. All of these forces shaped the current state of the market and will effect the future directions that the molecular modeling community takes.

The growing maturity of the industry

At the beginning of 1990, over 25 independent companies were part of the computational chemistry community; and the industry continued to expand during the 1990s with the addition of:

- Molecular Applications Group (1990, Michael Levitt and Chris Lee), a molecular modeling and protein design company,
- Schrödinger (1990, Richard Friesner, William Goddard and Murco Ringnalda), which marketed PS-GVB, an *ab initio* based quantum chemistry program,
- Wavefunction (1991, Warren Hehre), whose primary product was the Spartan *ab initio* suite of programs,
- Synopsys (1992, Glen Hopkinson, Keith Harrington and Paul Hoyle), whose Accord series of chemical information tools quickly caught the attention of the industry,
- SemiChem (1992, Andrew Holder), which marketed the PC version of AMPAC,
- Q-Chem (1993, Carlos Gonzalez, Peter Gill and Benny Johnson), whose product was the *ab initio* quantum chemistry software, Q-CHEM,
- Interactive Simulations (1994, Mark Surles), whose Sculpt program provided desktop protein modeling and display,
- ChemSoft (later named SciVision, 1994, Joseph Votano),
- Biosoft (later named MolSoft, 1994, Ruben and Margarita Abagyan),

1359-6446/06/$ - see front matter doi:10.1016/j.drudis.2008.04.008

- Chemical Computing Group (1994),
- SimBioSys (1996),
- Molecular Networks GmbH (1997, Johann Gasteiger),
- OpenEye Scientific Software (1997, Anthony Nicholls),
- LION Bioscience AG (1997, Friedrich von Bohlen and Thure Etzold),
- BioReason (1998, Susan Bassett, John Elling and Tony Rippo) and
- Scitegic (1999, Matthew Hahn and David Rogers).

Although some of these start-ups would either remain small or disappear altogether over the next 17 years, several would play major parts in the dynamics of the market during this period.

By 1990, the computational chemistry marketplace had segmented into three distinct strata: the industry leaders in their respective areas (Biosym, Polygen and Tripos), a group of growing companies led by BioDesign, Chemical Design and Oxford Molecular, and several smaller companies that provided specialized products. The problem that the larger companies faced was one of market saturation. As the use of computational chemistry was growing, the number of groups within the pharmaceutical industry remained limited and the size of these groups was small compared to other departments within the industry. When coupled with the specialized nature of each company's products, the immediate impact of this was that, over a relatively short period of time, the software vendors had reached nearly all of their potential customers. In addition, all of the companies found that supporting their current software and continuing to add the new science demanded by an increasingly sophisticated customer base consumed more resources than they had envisioned and that the cost of this support was impossible to cover with the annual fees charged for software maintenance. As a result, the vendors began to explore alternatives to add to their market share including the introduction of new science; the spread of their existing platform to new groups of customers, and the acquisition of their competitor's market.

The increasing number of molecular modeling products

One of the first application areas where companies began to release competing products was at the interface of small-molecule modeling and macromolecular analysis. Tripos, a company whose strengths were on the basis of small-molecule design (pharmacophore analysis, conformational searching, QSAR, etc.) introduced a biopolymer module that included homology techniques from Blundell's research [8] and code for molecular dynamics. On the west coast, Biosym, a macromolecular modeling-based company, moved into the small-molecule analysis/QSAR arena by aligning with Valery Golendar and marketing his APEX program [9–11]. Although these companies anticipated that allowing scientists to easily move between small- and large-molecule capabilities within the framework of their software would help to secure their customer base and provide an entry to those groups where they did not have a presence, they did not anticipate that the availability of software with nearly identical functionality would also lengthen the sales cycle because companies now performed rigorous side-by-side comparisons of their products.

As the molecular modeling companies began to expand their products into areas addressed by their competitors, they also began to see a major shift in the pharmaceutical industry's approach to research. Before the mid-1980s, the traditional approach to compound discovery was on the basis of the analysis of biological pathways and the development of screening systems that could detect new chemical entities that altered responses within these pathways. During the mid- to late-1980s, the adoption of combinatorial approaches to compound discovery completely changed this research paradigm. As a result, molecular modeling companies were forced to move their development efforts to tools to assess molecular diversity, design combinatorial libraries, create screening sets based on molecular similarity and, eventually, automate the docking of large compound libraries into receptor models. Several of the larger computational chemistry companies introduced these products during the 1990s, while others used the gap between products and needs to start new companies on the basis of this new approach to research.

Some of the products that addressed this shift in research began to appear in the early 1990s. In addition to simulation techniques, computational chemistry companies began to hear from their customer base that new approaches to storing and searching structures, especially 3D structures, would assist their research efforts. One company, BioCAD, was founded specifically to address the 3D searching problem, while Daylight, a company founded as an outgrowth of the Pomona MedChem project, was created to use an entirely novel approach to the enterprise of chemical registration, searching and analysis. By contrast, the molecular modeling organizations undertook database projects in an effort to increase the spread of their software into departments beyond computational chemistry and to move this business away from MDL, which offered MACCS-3D (and later ISIS-3D) [12,13]. Tripos initiated an industry-based consortium to develop the Unity 3D database product [14], Chemical Design marketed its Chem-3DB [15] add-on to Chem-X and Polygen announced an information management project that was known as Centrum [16]. Biosym, rather than developing competitive database products, formed an alliance with MDL to address this need. Although the new products provided the added functionality of performing searches on 3D structures, many of the products were also developed to offer an alternative to the vendors' perception that the MDL monopoly was ready to be broken. As the modeling vendors discovered, however, MDL's products were so deeply entrenched in their customer's IT infrastructure that it was cost-prohibitive to replace them. As a result, the database revenues that were expected at many of the modeling companies did not materialize. However, these developments did highlight three important issues that were to impact discovery research as well as the products that would be developed by the molecular modeling vendors:

- The growing influence that computational chemistry was exerting within research programs,
- The beginning of an appreciation for the importance that data flow management should play in research and development,
- The problems of data incompatibility between programs and therefore the amount of time spent converting file formats so that data could be used by research scientists.

In addition to database projects, many of the modeling companies were also seeking to develop computational products that supplemented their existing software. The results of these efforts included an industry-wide move into polymers. Chemical Design

and BioDesign brought out materials modules and Biosym began an exploration of materials and semiconductor simulations. The vendors also attempted to move from the small group of computational chemists into a larger market – the bench chemist. In addition to the introduction of Chem3D (CambridgeSoft), CAChe, SemiChem, HyperChem and other PC-based modeling packages, the molecular modeling vendors attempted to streamline their existing products by simplifying the interface. Tripos utilized their SYBYL Programming Language to introduce the LabVision version of SYBYL [17] and Chemical Design ported Chem-X to the PC and Macintosh computers.

Although a significant amount of time and money was expended on these ventures, none of them paid off for the industry in the long run. The only product that made any inroads into medicinal chemistry/organic chemistry was the MacroModel software from Clark Still's research group at Columbia University. In the polymer/material software effort, the chemical industry suffered a major downturn in the midst of the vendor's development cycle, which created a significant cash shortfall for the modeling companies that had invested in technology development for this market.

The impact of the financial analysts: market consolidation

As the 1990s progressed, it appeared that there was 'explosive' growth in the number of computational chemistry companies and continuing increases in the size of computational groups within the pharmaceutical industry. Reports in the popular press thus piqued the interest of companies outside the traditional science community and financial analysts began to take note of the promise of computational chemistry and its potential impact on reducing the cost of research. This interest started in 1989 with the publication of a market research study, *Conflicting Trends in Computational Chemistry* [18], and peaked with the follow-on 1991 report *Growth and Opportunity in Computational Chemistry* [19], both reports were published by The Aberdeen Group. In addition to profiling the current molecular modeling hardware and software companies, the report analyzed the business and concluded that the future growth of the sector would be 35% per year and that the industry group would reach a total revenue of over US$ 1.4 billion by 1994. The 1991 report summarized the financial community's enthusiasm for simulation (Box 1).

Based in large part on this analysis of the industry, several organizations began the rush to capitalize on the perceived leverage that the computational chemistry industry represented.

The first indication that major corporations were about to enter the marketplace was the investment position that IBM made in Polygen. This was followed in 1992 by the exclusive license that Autodesk signed for the right to market Hypercube's products. Despite, or perhaps because of, the increased interest within the financial community, the start of the wave of computational chemistry company mergers also began in 1991. In many cases these mergers were necessary to maintain a competitive edge in the marketplace. In other cases, organizations looking to capture the current marketing 'sizzle' added the computational component to their research efforts.

BioDesign, which was focused on the life sciences market, began to move into polymers (POLYGRAF) and materials (XTALGRAF). To reflect this shift, the company changed its name to Molecular Simulations, Incorporated (MSI) in 1989 and moved its office to the San Francisco area. In 1990, it acquired Cambridge Molecular Design (and their Cerius suite of materials design and simulation software) and, early in 1991, merged with Polygen to form the new MSI; the new company retained dual headquarters in Sunnyvale and Cambridge. During 1992, Biosym acquired Hare Research and was itself acquired by Corning. In addition, Maxwell Communications purchased ORAC and merged it with MDL and Evans & Sutherland purchased New Methods Research (NMRi, the NMR-based data acquisition and analysis company created by George Levy during 1983) and merged it with Tripos. Then, in 1993, Oxford Molecular established its operational strategy when it made its first acquisition, Biostructure. Over the next seven years, Oxford would acquire several companies including MacVector from Eastman Kodak, UniChem from Cray Research, Intelligenetics from Amoco Technologies, Health Design, Genetics Computer Group, MLR Automation and Chemical Design.

Interest in the computational chemistry marketplace may have reached its height when, in 1993, MDL filed an IPO as MDL Information Systems at a valuation of US$ 90 million. The next year, Evans & Sutherland followed with the spinout of Tripos Associates as a public company through an offering that was essentially a stock split to current E&S shareholders and Oxford Molecular Group released their IPO on the London Stock Exchange. During that year, MSI continued its growth through mergers and acquisitions with the purchase of BioCAD.

The public offerings of MDL and Tripos were a major step for the molecular modeling industry in that this offered the first opportunity for the public disclosure of the income and expenses for a computational chemistry-based company. In their 1994 annual report, Tripos reported sales of US$ 19,602,000 and a net income of US$ 352,000. MDL reported revenues of US$ 43,381,000 and a net income of US$ 37,000. Given that these were considered to be among the industry leaders, the financial projections published for this industry were suspect. Assuming that Biosym, MSI, Oxford Molecular and Chemical Design had revenues similar to Tripos', it is probable that total sales for the computational chemistry market did not exceed US$ 140 million to US$ 150 million in 1994 and that profit margins were very slim. Two years later, the revenues for the industry, based on public reports for the four largest companies, were still less than US$ 200 million.

By 1995, competing events highlighted a splintered opinion of the computational chemistry businesses. Oxford Molecular continued to expand its customer base by acquiring CAChe Scientific as well as the RS^3 product line from PSI. Corning decided that two companies competing for protein simulation customers was not viable and so they purchased MSI and combined them with

BOX 1

'The world's chemical and allied industries annually invest hundreds of millions of dollars, deutschemarks, francs, yen, and other currencies in an attempt to discover and develop new products using time-consuming, labor-intensive, repetitious methods little changed from the 19th Century. That is the nature of chemical discovery. If and when computers can reduce those efforts by even a few percentage points, the payoff will be measured in hundreds of millions and even billions of dollars.'

Biosym to form an independent company, MSI, with headquarters in California. On the public front, Chemical Design began the process of moving to list the company on the London Exchange, which was completed in 1996. In contrast to these events, Autodesk quietly returned exclusive rights for the Hyperchem product line to Hypercube and left the computational chemistry market.

The last part of the 1990s continued to see an increase in mergers and acquisitions as companies worked to increase their valuations. Major moves were made by MDL Information Systems, which agreed to be acquired by Reed-Elsevier for US$ 320 million, and Molecular Simulations, which was acquired by Pharmacopeia. The MSI purchase was financed through a tax-free stock-for-stock exchange of 8.5 million newly issued shares of Pharmacopeia common stock.

Reed-Elsevier stated that MDL's customers would find value in the combined entity and the MDL software would be a natural delivery vehicle for Elsevier's content, which included the Beilstein database (acquired in 1998). During the 1990s, MDL also acquired Interactive Simulations, Afferent Systems, a combinatorial library management company and SciVision. With the acquisition of MSI, Pharmacopeia stated that the molecular diversity tools being developed by MSI would provide a competitive advantage in the contract services arena. Neither company anticipated the suspicion that was generated within the user community by these mergers. The final two changes that occurred in the industry at the close of the decade were the acquisition of the MacroModel program by Schrödinger and the purchase of Molecular Application Group's assets by Celera.

The 1990s also marked a change in the software model for the computational chemistry industry. In 1994, Paul Labute, George Lowenfeld and Jose Cabral founded Chemical Computing Group (CCG). Rather than marketing a monolithic modeling package with components dedicated to specific computational tasks, CCG provided the Molecular Operating Environment (MOE) and the Scientific Vector Language (SVL) that allowed scientists to implement their own functionality within a development environment that was specific for molecular modeling. As the first of the 'toolkit' approaches to modeling, the software from CCG created a significant level of interest in the community. In addition, the company priced their software well below the cost for other molecular modeling systems, thus removing one of the significant barriers to the entry to this market.

In addition to CCG, two other companies that offered an alternative approach to product development and marketing were Schrödinger and OpenEye Scientific Software. Both of these organizations utilized scientists from the pharmaceutical industry to assist in the definition and refinement of their products. And, like CCG, they offered tools in response to specific problems rather than a full-featured platform product. In addition, the organizations rediscovered the 'family' concept of user groups that was a part of the early culture of the molecular modeling community (i.e. when you purchase products from the company and join the user group you become part of the family that is dedicated to advancing the science of simulation).

As the new decade began, the consolidation wave that started during the 1990s showed no signs of slowing down. Pharmacopeia added Synopsys and Oxford Molecular (minus the CAChe component, which Oxford sold to Fujitsu earlier in the year) to their Molecular Simulations purchase and named the combined organizations Accelrys. Accelrys initially offered many of the products that were available from the original companies; however, by the time it was spun out of Pharmacopeia as an independent company in 2004, the product choices were drastically reduced, primarily to those originally marketed by MSI. By 2007, several of the original programs were folded into Discovery Studio and Materials Studio or were sold to new organizations such as MacVector (formed in 2007 to support the Macintosh bioinformatics programs) and FelixNMR (formed in 2007 by Steve Unger to continue the development of Felix). Accelrys also continued to pursue their philosophy of growth through M&A and, in 2004, the company acquired Scitegic, a data integration company formed in 1999 by former MSI employees, Matthew Hahn and David Rogers. This acquisition, at a cost of US$ 21.5 million, marked another turning point in the industry – the shift in focus from pure computational software to products that focused on data flow, integration and management.

Conclusion

Since this retrospective on the evolution of the molecular modeling industry began with the creation of the industry's first two companies, it is perhaps fitting that it ends with them as well. In 2007, Elsevier sold the MDL organization to Symyx for US$ 123 million. As a part of the acquisition, Elsevier retained the scientific-content-related assets of MDL, including the CrossFire Beilstein, PharmaPendium and Patent Chemistry databases. In a move similar to Accelrys' with its acquisition of Scitegic, Symyx announced plans to integrate the MDL Isentris products with Symyx's research services and software for electronic laboratory notebooks, lab instrumentation systems, lab automation and scientific decision support. The valuation of MDL over a period of 13 years indicated the speculative nature of the computational chemistry marketplace. In 1994, the company issued an IPO at US$ 90 million based on revenues of US$ 43 million. Three years later Elsevier valued the company at US$ 320 million, and ten years after this acquisition the company was valued at just US$ 123 million.

At Tripos, events were even less kind. After several years of searching for the business model that would lead to sustained profitability, CEO John McAlister presided over the sale of the company's assets and the original organization was dissolved. The informatics portion of the company, with 2006 revenues of US$ 27.4 million, was purchased for a mere US$ 26.2 million by a corporate turn-around specialist, Vector Capital, while the discovery services organization, Tripos Receptor Research, with 2005 revenues of US$ 27.4 million, was purchased by Commonwealth Biotechnologies for US$ 2.15 million and, in late 2007, was renamed Exelgen, Ltd. As with MDL, the impact of these events within the user community is still being evaluated.

The dynamic nature of this industry group was reflected by the fact that, at the beginning of 1980, two organizations were identified as computational chemistry companies, but by the beginning of 2000, 34 companies were identified as modeling and simulations organizations. And this growth occurred despite the large number of mergers and acquisitions over the 20-year period. In addition, the industry group witnessed several changes in the mix of products that their customers required.

The initial excitement of molecular modeling, the ability to display and manipulate chemical structures on a computer terminal, morphed into an emphasis on simulations, then to the need to analyze large datasets, and finally into the collection and management of data. Given that the market was driven by the needs of the research programs within the chemical industries, it is not surprising that the molecular modeling marketplace was a dynamic business. When an existing company either folded or was purchased a start-up with a new approach to the business of simulations would emerge to fill an underserved niche. If one looks at the similarity to other business segments, there are several similarities with the growth of Internet-based technologies; products grow and evolve in an environment where the basis of product features is in constant flex. And, with the pharmaceutical industry currently poised to change in response to market forces within its own business segment, it can be anticipated that the molecular modeling industry will continue to evolve as well.

References

1 Visiers, I. (2007) Principles of G-protein coupled receptor modeling for drug discovery. In *Annual Reports in Computational Chemistry* (Spellmeyer, D.C. and Wheeler, R.A., eds), . pp. 209–227, Elsevier

2 Egan, W.J. (2007) Computational models for ADME. In *Annual Reports in Medicinal Chemistry*, (Vol. 42) (Macor, J.E., ed.), , Vol. 42. pp. 449–467, Elsevier

3 Güner, O.F. (2002) History and evolution of the pharmacophore concept in computer-aided drug design. *Curr. Top. Med. Chem.* 2, 1321–1332

4 Kubinyi, H., Folkers, G., Martin, Y.C., eds (1998) *3D QSAR in Drug Design: Recent Advances*, . Kluwer/Escom

5 Gupta, M.K., Gupta, S.P., Hadjipavlou-Litina, D., Hannongbua, S., eds (2006) *QSAR and Molecular Modeling Studies in Heterocyclic Drugs II*, . Springer

6 Nikolova, N. and Jaworska, J. (2004) Approaches to measure chemical similarity – a review. *QSAR Comb. Sci.* 22, 1006–1026

7 Lipkowitz, K.B. and Boyd , D.B., eds (1990–2007) Reviews in Computational Chemistry, Vol. 1–27. VCH Publishers

8 Blundell, T.L. *et al.* (1978) Insulin-like growth factor: a model for tertiary structure accounting for immunoreactivity and receptor binding. *Proc. Natl. Acad. Sci. U. S. A.* 75, 180–184

9 Avidon, V.V. *et al.* (1982) Structure–activity relationship oriented languages for chemical structure representation. *J. Chem. Inf. Comp. Sci.* 22, 207–214

10 Golender, V.E. and Rozenblit, A.B. (1983) *Logical and Combinatorial Algorithms in Drug Design*. Research Studies Press

11 Golender, V.E. and Vorpagel, E.R. (1993) Computer-assisted pharmacophore identification. In *3D QSAR in Drug Design: Theory, Methods and Applications* (Kubinyi, H., ed.), . pp. 137–149, ESCOM

12 Christie, B.D. *et al.* (1990) Database structure and searching in MACCS-3D. *Tetrahedron Comput. Methodol.* 3-6C, 653–664

13 Moock, T.E. *et al.* (1994) Conformational searching in ISIS/3D databases. *J. Chem. Inf. Comput. Sci.* 34, 184–189

14 Hurst, T. (1994) Flexible 3D searching: the directed tweak technique. *J. Chem. Inf. Comput. Sci.* 34, 190–196

15 Murrall, N.W. and Davies, E.K. (1990) Conformational freedom in 3-D databases. 1. Techniques. *J. Chem. Inf. Comput. Sci.* 30, 312–316

16 Centrum was never released as a product, but the concepts of this project have been implemented in several generations of project/document management software

17 LabVision was never released as a formal product, however, the ideas behind it formed the basis for a follow-on product, Benchware, a suite of Microsoft Windows-based applications offering chemical library enumeration, design and visual exploration, that was introduced in 2004

18 *Conflicting Trends in Computational Chemistry* (1989). The Aberdeen Group, Boston, MA.

19 *Growth and Opportunity in Computational Chemistry* (1991). The Aberdeen Group, Boston, MA.