The optimal fragmentation principle – Reply

Initial letter: Johnson, D.E. (2001) The optimal fragmentation principle. *Drug Discovery Today* 6, 175

Response from Peter H. Bach

Let's make a full meal of IT!

Why only the entrée? Man deserves nutritious meals, perhaps several courses (efficacy, toxicity and metabolism prediction), with the right flavours, using the highest-quality ingredients (would mother want anything less!) from all over the world.

In addition to large pharmaceutical companies providing key inputs, there are other sources of drug development 'ingredients' that could make *in silico* toxicology (IST) work successfully. The regulatory and health agencies have full safety packages in the context of human use and the experience of seeing a big picture. Similarly, the contract research organizations (CROs) have the potential to provide information on many molecules that have failed to go into development because of toxicity. There are already precedents for collaboration between stakeholders who prepare, serve and consume material on which predictive informatics can be based:

- 'DEREK' (Deductive Estimation of Risk from Existing Knowledge) the LASHA (LHASA Ltd, University of Leeds, Leeds, UK; http://www.chem.leeds.ac.uk/luk/) expert system involves corporate members who help address the interests of the companies concerned, based on in-house safety data.
- The Food and Drug Administration (FDA) has shared confidential data in the past with MultiCASE (Beachwood, OH, USA; http://www.multicase.com/).

In addition, the academic community and public domain material has important information and, more relevantly, provides mechanistic insight into pathology lesions. This should enable an additional understanding of where the different processes that lead to the same final toxic end-points (pathology lesion) can affect the accuracy of the prediction.

One burning question for drug developers is 'where should we eat?' Each IST restaurant has a different basis for prediction, which must impart unique strengths and weaknesses. Few individuals would commit themselves to only one restaurant. There should be better ways to access all of the options and select the most appropriate option from any one cuisine when required. The IST industry must find a common platform that will link all the different options (in a user-friendly way) so that rational choices can be made. One possibility is a 'pay as you go' approach (already being offered by i-Tox.com at http://www.i-Tox.com), an associate company of MultiCASE that offers an Internet-enabled *in silico* search system on the basis of pre-paid access. This would increase the use of informatic portals by small companies, who currently can be locked out by expensive licensing agreements. Pay as you IST could also offer the potential for the many small players to add information to the fragment database(s) from their own compounds.

That is all very nice in theory, but chefs rarely part with their best recipes. Industry, CROs and regulatory agencies are bound by commercial interests and confidentiality. The most important challenge is how to convince the *Supremo Chef de Cuisine* (CEOs, Directors, lawyers and accountants) at each stakeholder that sharing information is good in terms of getting safer medicines to the market more rapidly. Corporate moguls (being delicate and far-sighted folk) are likely to be concerned that 'chemical-fragment fruit-salad' could be reassembled to whole fruits, with loss of competitive advantage: not easily reconciled with sharing or common good. Indigestion is the last thing that shareholders would want if their own chemical fragments were used by a competitor to produce a lemon or tangelo (ugli-fruit) for them!

There are many difficult questions for which a consensus is required. The most important question is who would ensure the contents of the Holy Grail on safety are kept safe. Is there a role for a group such as the ICH (International Committee on Harmonisation, Geneva, Switzerland) or ILSI (International Life Sciences Institute, WA, USA) to oversee such a sensitive, but vital development? Will there be a large enough buy-in to make it work? Where does ownership of the different aspects of such an approach start and stop?

The one thing that is more certain is that we will all have to eat out of the same pot. Those who cannot stand the heat in the IST kitchen will not be there when the *pièce de résistance* is served; and it will be served! All of the stakeholders need to start cooking better and safer drugs (rather than just talking about ingredients and flavours) and thinking of swapping recipes. The technology is there to do a good job; the missing ingredient should not be the will to make it happen.

If we do not finish everything on our 'IST plate', mother may decide we should not get dessert, which would be our just deserts for not having tried harder.

Peter H. Bach
Inveresk Research, Tranent
UK H33 2NE

The optimal fragmentation principle – Reply ▲

Initial letter: Johnson, D.E. (2001) The optimal fragmentation principle. *Drug Discovery Today* 6, 175

Response from Ann Richard

To follow up on my comment to Dale Johnson and to respond to his entertaining and provocative letter, I confess to some frustration in seeing a show-and-tell of a sophisticated tool for

1359-6446/01/$ – see front matter PII: S1359-6446(01)01716-0

data mining and exploration, without seeing how the tool will realistically benefit the research community working in the toxicity modeling area. The fact is that these new and costly data-mining tools require large stores of data relative to a defined endpoint, and are being marketed to, and will primarily benefit, the data-richest among us – i.e. the pharmaceutical and chemical industries.

Lack of information

Despite the many common factors between approaches, structure-based toxicity prediction historically has been a poor cousin of drug design. As opposed to highly focused efforts to optimize drugs for particular and often well-understood biomolecular interactions, toxicity prediction attempts to anticipate one or more potential adverse effects. These effects can occur by a variety of mechanisms and there is often little understanding of, and a lack of data relating to, many of these mechanisms.

It is widely acknowledged that access to quality toxicity data for a sufficient number and diversity of chemicals across the spectrum of toxicity endpoints of potential interest is the major factor limiting development of useful structure–activity models and expert toxicity prediction systems. The fact is, the larger and more diverse the training set, the better the toxicity prediction model.

To move forward, we must accept that proprietary toxicity information coupled with full knowledge of the chemical structure would reveal too much about drug development leads and strategies to ever see 'the light of day'. Unfortunately, designed biological activity and undesired toxicity are often chemically entangled. So where does this leave us? As the great basketball coach, John Wooden, offered, 'Do not let what you cannot do interfere with what you can do.' This brings us to where Dale's letter left off, asking what practical measures we can take, across the public, private/corporate and commercial domains, to make substantive gains in furthering toxicity modeling capabilities. To follow up on Dale's alliterative thesis pertaining to the Optimal Fragmentation Principle, I offer below some comments and pose some modest challenges to the three corners of the triangle in this struggle.

Public efforts

In the public domain, there are several efforts currently underway to consolidate toxicity data by coupling it with chemical structure and property information, and to make these data more widely available and useful for toxicity modeling efforts. Examples consist of web-based tools (e.g. http://toxnet.nlm.nih.gov/, http://chemfinder.camsoft.com/, http://cactvs.cit.nih.gov/ncidb/), a proposed distributed common-format toxicity database paradigm, and an international non-profit toxicity database industry consortium (organized by the International Life Sciences Institute, Washington DC, MD, USA). These measures offer the promise of lifting toxicity modeling to a new plane of scientific participation and involvement, fostering greater collaboration between the toxicology, chemistry and modeling communities and providing more effective linkages between toxicity and chemical information spanning multiple biological endpoints. It should be acknowledged that these largely public efforts will provide a valuable service to both commercial and corporate concerns interested in improved access to toxicity information and improved toxicity models. Hence, as we all have something to gain from these efforts, we all should support and participate in these efforts to make toxicity data more publicly accessible and usable.

Corporate efforts

For the pharmaceutical and chemical industries, it is important to acknowledge the considerable economic incentives to improved toxicity prediction models. There is also a need for community-wide pooling of toxicity-related information and knowledge if we are to make significant headway with this problem. There are two essential corollaries. The first is that creative technologies for extracting useful pieces of toxicity information and knowledge from corporate databases, while shielding the identity of proprietary chemical structures, must be developed. The second is that the modeling community must be able to effectively utilize this extracted information to improve existing models without having access to proprietary chemical identities.

I am aware of at least a few examples of successful cooperative ventures in which these goals are being met. One example is between a chemistry modeling vendor of physicochemical properties and a chemical industry company (reparametrizing the physicochemistry predictor using in-house industry data stores). A second example is provided by the DEREK user consortium (LHASA Ltd, Leeds, UK) that contributes generalized rules to toxicity prediction rulebases. A third example is a cooperative venture between MultiCASE (Beachwood, OH, USA) and the US Food and Drug Administration (FDA) to develop improved fragment-based toxicity prediction for pharmaceuticals. In all three cases, the available public data are used to validate the newly parametrized models or provide independent support for newly constructed rules. If there is a will, there is a way, but industry must see that these efforts best serve their longer-term interests and be willing to engage in such efforts.

In addition, if we can find ways to bring these cooperative efforts to pool data and knowledge into the public domain (perhaps managed by a non-profit or government website) we can encourage even greater scientific participation and creative efforts for developing useful models.

1359-6446/01/$ – see front matter PII: S1359-6446(01)01717-2

Commercial efforts

To the Dale Johnsons of the world that develop and market models and data-mining tools that could potentially be used for either toxicity prediction or drug/chemical design, I appeal to you to consider that the toxicity prediction problem most appropriately resides in the public domain, yet deserves the creative involvement of the commercial sector.

Vendors have an important role and motivation to develop technologies that effectively shield but extract useful toxicity information from individual clients to better serve the larger potential client community. In addition, as more public toxicity databases become available in standard structure-searchable format, vendors as well as the public will have easier access to toxicity data to demonstrate the use of commercial tools on a public forum, such as a website. For example, it is becoming common practice for vendors to provide services and resources on their websites, as a draw for potential clients and as a service for current clients. Even beyond a mere demonstration, however, vendors could offer a real and valuable service by providing functional enhancements to the public toxicity databases. The commercial opportunities would be in advertising their wares and in marketing the same tools for exploration of public/private corporate databases for all types of biological activities (e.g. ADME properties or pharmaceutical activities). It seems that it could be a clear win–win situation.

Mutual contributions

Pharmaceutical and chemical companies are understandably eager to obtain greater access to public toxicity data and to take advantage of any effort that brings toxicity data and improved toxicity prediction models into the public or commercial arena. But are they willing to accept that they can and should contribute to these efforts, and that it is in their best, long-term interest to do so? What has to be avoided is a narrow corporate mindset that views knowledge mined from corporately held toxicity data sources, whether public or private, as a proprietary resource conferring a competitive advantage in the marketplace.

In closing, and in honor of the man to whom this particular day that I write is dedicated, I quote Martin Luther King, Jr: 'Every man must decide whether he will walk in the light of creative altruism or in the darkness of destructive selfishness.' Perhaps through creative means and determination we, as a community of public, corporate and commercial concerns, can choose the best, most well-lit path for tackling the formidable problem of toxicity prediction.

Ann M. Richard
National Health and Environmental Effects Research Laboratory
US Environmental Protection Agency
Research Triangle Park, NC, USA