# Opening the 'Black Box': Interpreting in Silico Models to Guide Compound Design

Matthew Segall, Ed Champness

**BioFocus DPI (Inpharmatica), 127 Cambridge Science Park, Milton Road, Cambridge, CB6 3WJ, m.segall@inpharmatica.co.uk**

## Introduction

*In silico* predictive models are now widely used to predict a range of molecular properties and help prioritise molecule for synthesis [1]. However, a common criticism often levelled at predictive models is that they offer few clues regarding **why** a molecule is predicted to have a certain property. By definition, models encode relationships between molecular structure and properties, but interpreting and visualising this information to design better molecules has been almost impossible. This is particularly true of models built with modern 'machine learning' techniques such as artificial neural networks (ANN), Gaussian processes (GP) or support-vector machines (SVM). The models that these techniques create have commonly been described as 'black box.'

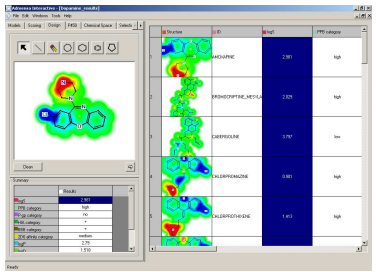

**Figure 1.** An example of the 'Glowing Molecule' within the interactive designer of Admensa Interactive™.

In this poster we will present a novel algorithm that 'opens' these black box models, providing an intuitive visualisation of these structural relationships. The objectives of this approach are to:

- Clearly identify 'problem' regions on a molecule
- Highlight functional groups that tend to improve a molecular property
- Understand the 'multidimensional' structure activity relationships (SAR) of a chemical series

The output of the analysis is a coloured field on which the 2D molecular structure is superimposed. For this reason, we call this the '**Glowing Molecule**.'

To enable interactive exploration of chemistry, this algorithm has been integrated with the **Admensa Interactive™** environment for decision support within drug discovery, to enable real-time feedback from predictive models as modifications are made to a molecule's structure.

## Methods*

A predictive model is a mathematical function that relates a set of descriptors (*x1,x2,x3…*) that characterise a molecule to a value of the property being modelled (*y*). Common descriptors include; simple 1D descriptors such as molecular weight, atom counts, 2D descriptors such as molecular fragments, topological polar-surface area (TPSA), 3D descriptors that capture information about the shape of a molecule or whole-molecule properties such as logP. The mathematical function *f(x1,x2,x3…)* that correlates with the predicted property is typically fitted to a dataset of molecules with known property values, using a statistical or 'machine learning' technique such as partial least squares (PLS), ANN, GP or SVM.

The mathematical function *f(x1,x2,x3…)* can be considered to define a *hypersurface* in the space of descriptors. In the simplest, linear case, $f(x1,x2,x3\ldots)=c1\times x1+c2\times x2+c3\times x3\ldots$, this is simply a plane as illustrated in Figure 2(a). However, for non-linear models such as those created using more advanced techniques, this hypersurface can adopt more complex forms, such as that illustrated in Figure 2(b).

The act of making a prediction for a molecule can be represented as finding the 'height' (*y*) of this hypersurface for a given set of coordinates (*x1,x2,x3…*), as shown in Figure 2. However, the shape of this surface contains additional information, in particular the descriptors that contribute most to variation in the property value. This is also illustrated in Figure 2.

It is notable that these trends are constant for a linear model, they do not depend on the particular molecule for which the prediction is being made. This reflects the relatively straightforward interpretation of linear models, where the influence of descriptors is simply related to the magnitude and sign of the coefficients in the equation *(c1,c2,c3…)*. However, they may vary significantly between molecules for non-linear forms of *f*.

If the contributions of each atom in a molecule to each descriptor can also be quantified, this enables the contributions of each atom to the overall trend in the molecule's property to be calculated. This is most easily achieved for 1D and 2D descriptors, but may be generalised to more complex descriptors. These contributions can then be represented as a colour in a spectrum and plotted behind a 2D depiction of the molecular structure for visualisation.


**Figure 2.** Examples of 'hypersurfaces' represented by predicted models. (a) illustrates a linear models and (b) a non-linear model. The trends captured by these surfaces in each case for the specified molecules are shown by arrows.

## Interpretation

This method identifies the descriptors and hence the regions of a molecule that have the strongest influence on the property value for that specific molecule, i.e. the analysis is local to a specific point in descriptor space and captures only short-range trends at that point. This gives rise to two effects that should be considered when interpreting the results of this analysis

The same descriptor may have a different degree of influence for different molecules. Figure 3 illustrates a non-linear relationship between a descriptor (*x*) and property (*y*). For molecule A a modification that gives rise to a small change in the descriptor value will have a dramatic impact on the property value. On the other hand, a small change in the value of the same descriptor for molecule B will have little or no impact on the property value.

A large change in the structure of a molecule may 'jump' to a different region of descriptor space which may have a different relationship between the descriptors and property. In this case, the local information regarding the trends between descriptors and property may no-longer hold. This is illustrated in Figure 4. The trend for molecule C indicates that increasing descriptor *x* will lead to an increase in the property *y*. However, the large change between molecules C and D actually leads to a decrease in the property value due to the non-linearity in the relationship.

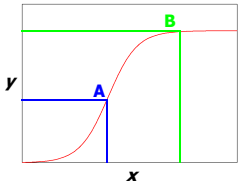


**Figure 3.** An illustrative 1-D example of a sigmoidal relationship between descriptor x and property y and the descriptor and corresponding property values for molecules **A** and **B**.

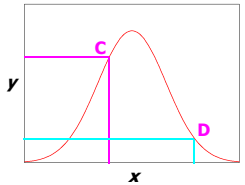


**Figure 4.** An illustrative 1-D example of a Gaussian relationship between descriptor *x* and property *y* and the descriptor and corresponding property values for molecules **C** and **D**.

## Illustrative Application

Rowley et al. [2] investigated a series of tryptamines with very high affinity and selectivity for the h5-HT2A receptor. Their main issue was to reduce affinity for the hERG IKr channel, measured by displacement of 4 nM [3H]-dofetilide binding to HEK cells stably expressing the hERG channel.

We have applied the 'Glowing Molecule' analysis to this series, using our proprietary QSAR model of hERG pIC50. As the model is based on 'gold standard' patch-clamp measurements, we expect only qualitative agreement with the dofetilide displacement results reported. However, Figure 5 demonstrates that the resulting visualisation provides good guidance on structural modifications.

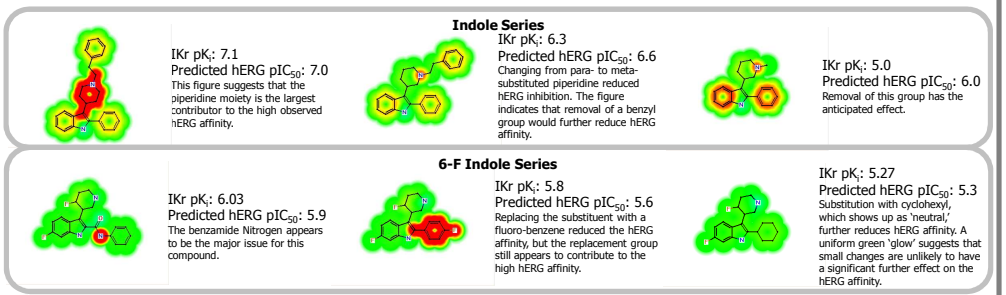


**Figure 5.** Examples of the 'Glowing Molecule' visualisation for a two series of hERG inhibitors.

## Conclusion

We have presented an innovative approach to providing additional interpretation of predictions produced by *in silico* models, opening the 'black box.' This approach provides qualitative guidance that highlights regions of a molecule that most strongly influence a predicted property. This can be used to interpret the SAR for a chemical series and help to guide redesign of compounds to overcome liabilities as illustrated by an example using published project data.

## References

[1] Segall et al. Exp. Opin. Drug Metab. Toxicol., 2006, **2**, pp. 325-337 [2] Rowley et al. J. Med. Chem., 2001, **44**, pp. 1603-1614

*Patent pending



Admensa Bringing balance to optimisation

BioFocus DPI A Galápagos Company