MedeA QT
An Interactive QSPR Toolbox

MedeA\textsuperscript{1} QT ("QSPR Toolbox") is an interactive toolbox that allows you to develop Quantitative Structure Property/Activity Relationship (QSPR/QSAR) models.

At-a-Glance

Employing MedeA structure lists as a convenient storage mechanism for collections of molecular information, MedeA QT enables you to construct correlative models, linking materials descriptors to properties and activities. With MedeA QT, you interactively adjust the details of QSPR models in order to describe any selected observation or activity for a given set of molecules or materials based on molecular descriptors, properties, and activities.

Key Benefits

- MedeA QT makes the construction of QSPR models efficient by allowing users to select optimal descriptor sets and to compare models easily
- Replaces ad hoc methods requiring manual data transformations and processing
- Reveals QSPR models, using molecular and structure based descriptors to predict a wide range of physical properties
- Interactively employs ordinary least squares (OLS) and partial least squares (PLS) methods to develop correlation models efficiently
- Provides insights into the key determinants and tradeoffs that govern complex phenomena, properties, activities, and process characteristics

As Cherkasov notes in a recent review\textsuperscript{2}, Corwin Hansch and colleagues have been long credited with developing correlative methods employed to extract insights from observed properties and descriptors for chemical systems. With MedeA QT, any desired observable or activity can be employed in creating QSPR models. The resulting QSPR models not only provide predicted properties for new systems, but direct insights into the relationship between physical properties and molecular or structural descriptors. When you use MedeA QT models, you can predict properties for materials that are not part of the training set. MedeA QT allows you to assess interactively the quality of the resulting QSPR models while varying the descriptors included in these models.

Any molecule or structure, or set of molecules, can be omitted from the training set employed in the development of the model. Models can then be immediately validated based on their ability to describe the properties of systems not included in the ordinary and partial least squares (OLS & PLS) procedures.

‘Hansch foresaw the power of large, searchable databases and put “inductive knowledge transfer” into practice’

Cherkasov et al., in ‘QSAR Modeling: Where have you been? Where are you going to?’

Once created, a MedeA QT model can be applied in the MedeA Environment to arbitrary structure lists with the same descriptors, either interactively or using MedeA Flowcharts. When used with MedeA Flowcharts, entire databases of materials can be efficiently annotated with computed model based properties.

MedeA QT makes the construction of QSPR models efficient, and eliminates the need to carry out ad hoc, manual data transformations and processing. MedeA QT is designed for use by experts and non-experts to gain insights into the origin of physical phenomena and to build links between observation and atomic descriptions.

Whenever sets of observations, even for properties that are challenging to simulate directly, can be obtained, MedeA QT can provide insights into the

\textsuperscript{1} MedeA and Materials Design are registered trademarks of Materials Design, Inc.
\textsuperscript{2} Cherkasov, A. et al., \textit{J Med Chem} \textbf{57}, 4977 (2014)
underlying causes of those varying properties. For example, the glass temperatures of epoxy ther-
mosets can be modeled based on experimental in-
formation, as obtained by Morrill et al.\(^3\).

The example of glass transition temperatures for
epoxy thermosets employs reported experimental
information collected by Morrill et al., and is an il-
lustration of the type of analysis that MedeA QT
facilitates.

Figure 1: Developing a transferable model for the heat of
vaporization using a training set containing 3,621 organic
molecules and a correlative model based on PLS.

Key Features

- Employs an efficient molecular spreadsheet
  system as its user interface
- Uses efficient partial least squares (PLS) based
  regression analysis to compute optimal
  regression models
- Any MedeA descriptor or property can be used
  as the basis for building QSPR models
- Stores MedeA QT QSPR models in an open
  and accessible XML based format, for
  straightforward inspection and publication
- MedeA QT QSPR models can be applied to
  MedeA structure lists interactively and to large
  scale structure lists using MedeA Flowcharts
- Highly interactive: automatic QSPR model
  updates, interactive descriptor selection and
  model comparison, interactive selection of
  training set members
- Correlates complex physical phenomena and
  their molecular and atomic level determinants

• Provides a foundation for more detailed
  atomistic simulations

\(^3\) Morrill Jason A., et al., Journal of chemical information and

Figure 2: Investigating the molecular determinants of the
glass transition temperature (\(T_g\)) of thermoset epoxy sys-
tems.

Required MedeA Modules

- MedeA Environment
- MedeA JobServer

Supported Modules

- MedeA QSPR
- MedeA P3C
- MedeA MOPAC
- MedeA GAUSSIAN

Find Out More

To learn more about how MedeA QT creates correlative models, contact
info@materialsdesign.com.

MedeA QT was developed in partnership with In-
tegrated Chemistry Design, Inc.