Enalos Nanoinformatics tools for the prediction of nanomaterials properties

NANOGENTOOLS EU Autumn School

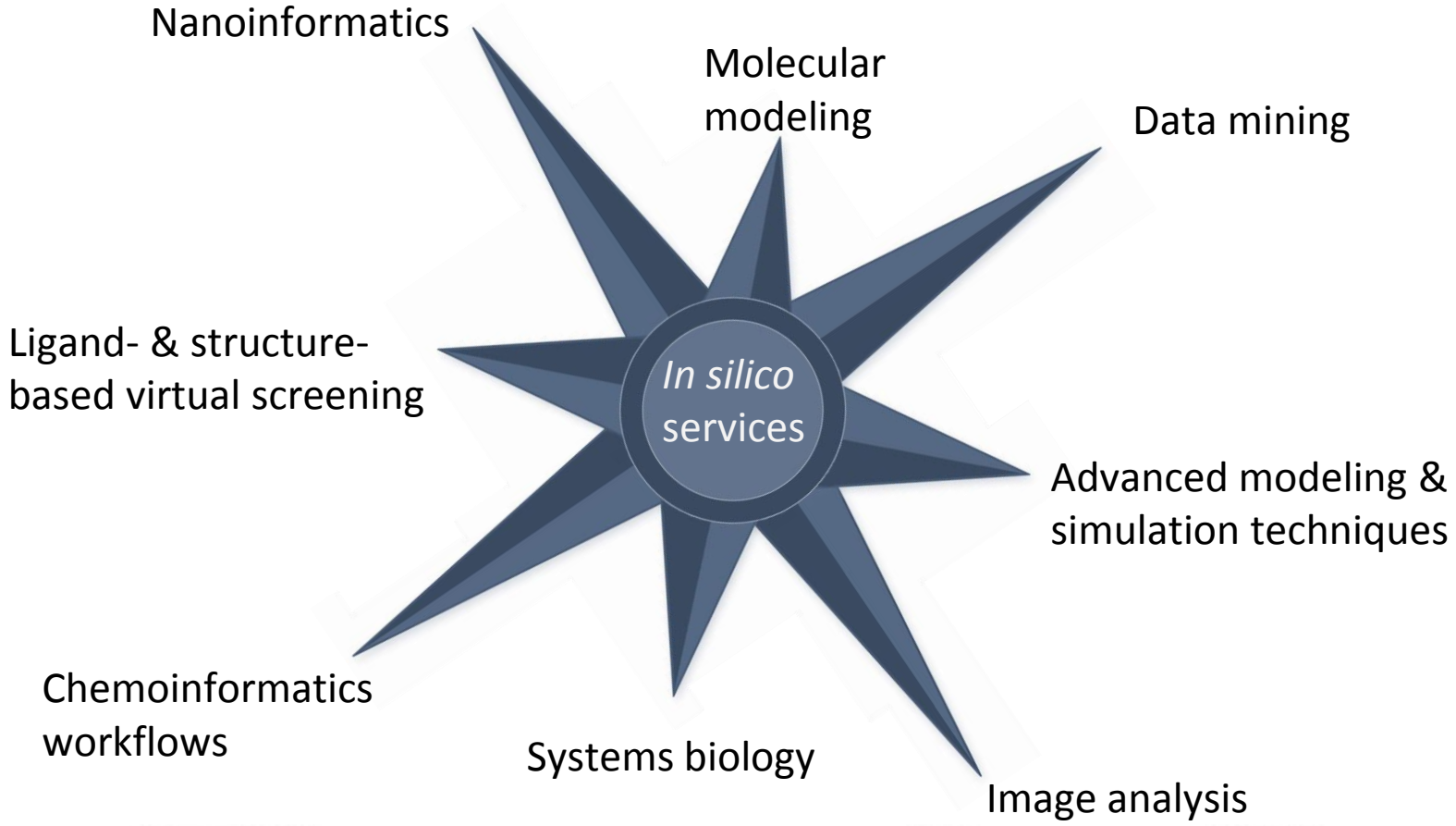
M. Eng. Dimitra Danai Varsou

Hotel Rice Palacio de los Blasones

Table of Contents

- Who we are
- Nanoinformatics
- Enalos+ software
- Enalos Cloud Platform for Nanoinformatics

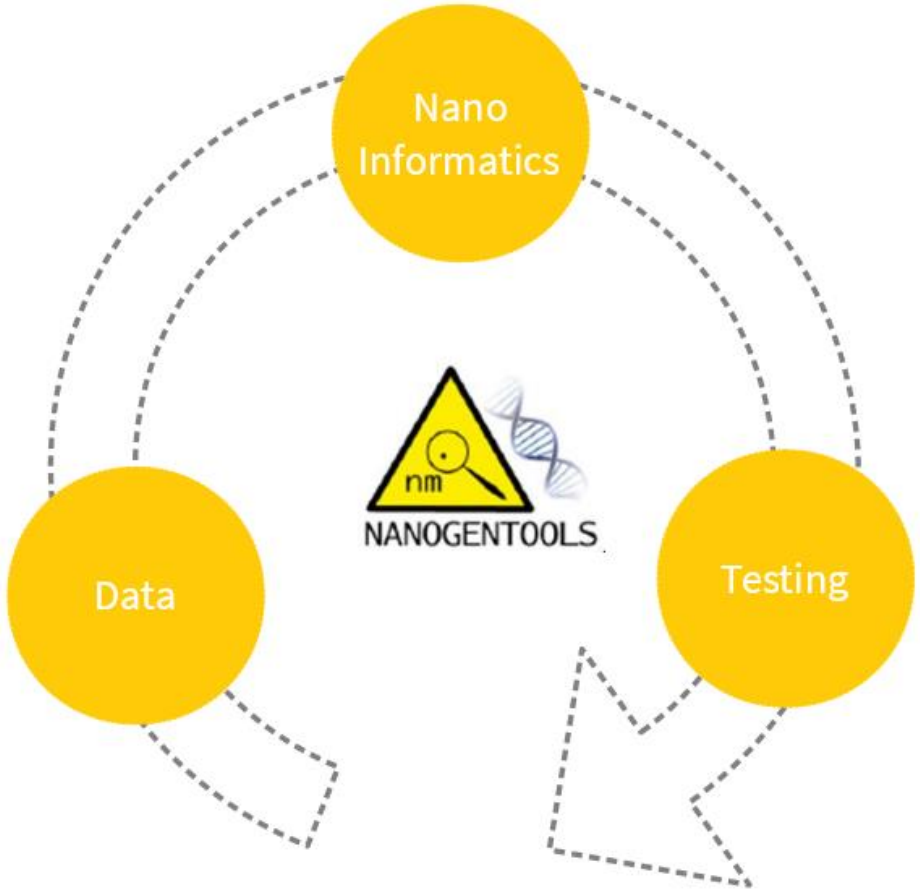
Work orientation



Investing in people

- According to Scopus NovaMechanics is the top Research SME in Cyprus
- All personnel is highly skilled with strong scientific background in the field of chemoinformatics, bioinformatics and medicinal chemistry
- Senior scientists have a strong academic record
- Managerial experience in large scale scientific projects, managed successfully EU & National funded projects

NovaMechanics in NANOGENTOOLS [1]



- In silico exploration of tested NMs
- Development of QNAR models
- Building risk assessment platform
- Prioritize NMs for biological evaluation
- Design of novel NMs with desired properties

NovaMechanics in NANOGENTOOLS [2]

- Meta-models
  - Meta-models development for the time demanding calculations of NMs quantum-mechanical (QM) and molecular dynamics (MD) simulations
  - Building a predictive modeling procedure to correlate all described input and output variables
  - The input/design variables will be selected among the QM and MD data and will be varied in a stepwise fashion to produce a large number of models
  - The outcome will be validated → robust and fast predictive models with well-defined domain of applicability for the prediction of QM and MD properties

NovaMechanics in NANOGENTOOLS [3]

- E-infrastructure/NovaMechanics server (key-feature: NVIDIA Tesla™ P100 12GB Passive GPU, 512GB RAM)
  - Speeding up the MD calculations procedures
  - Hosting GPU-accelerated databases
  - Streaming, processing, querying and analyzing datasets in seconds to milliseconds, instead of hours to minutes
  - GPU-parallelized processing architecture allows linear scalability and reduces analytical processing times for multi-billion row data sets
  - Application of time demanding state of the art modelling methodologies such as deep learning, in real time

Nanoinformatics

What is all about?

Development of a QNAR model

Risk assessment platform

What is all about? [1]

Main goal: Toxicity assessment of ENMs

Classical approach: *in vivo* and *in vitro* testing

Engineered NPs

Raman spectroscopy, TEM, FTIR, DLS, mass spectrometry, HTS, etc.

Toxicity endpoints:
cell viability, cell membrane damage, mitochondrial damage, DNA damage, genotoxicity etc.

What is all about? [2]

Main drawbacks

Time-consuming experiments

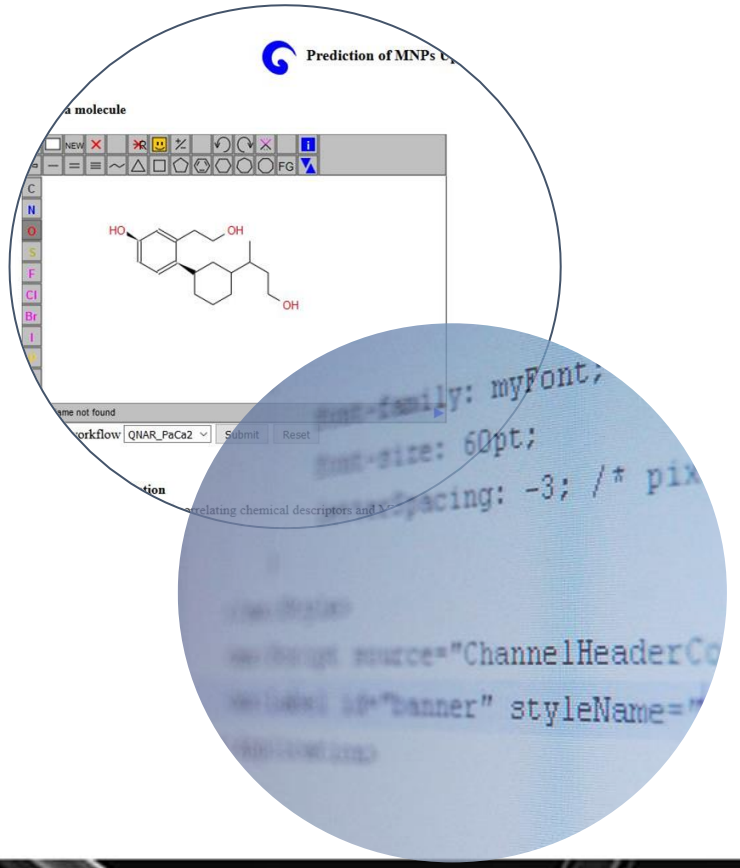
ENMs currently emerging in commercial applications

Expensive experiments

Use of laboratory animals

What is all about? [3]

- *In silico* testing:
  - Computational approach of the toxicity assessment of ENMs
  - High accuracy predictions of the potential toxic effects of ENMs
  - Development of user-friendly tools (web-services) for nanotoxicity assessment
  - Prioritization of ENMs for biological evaluation
  - Reduction of the time and the cost of experimental procedures



Development of a QNAR model [1]

Quantitative Nanostructure-Activity Relationship (QNAR) modelling

Properties of ENMs	Model	Prediction
Predictor variables (X) Physicochemical properties Structure Molecular descriptors	“Function” that relates the ENMs properties to their activity profile	Response variable (Y) Toxicity profile Endpoint value

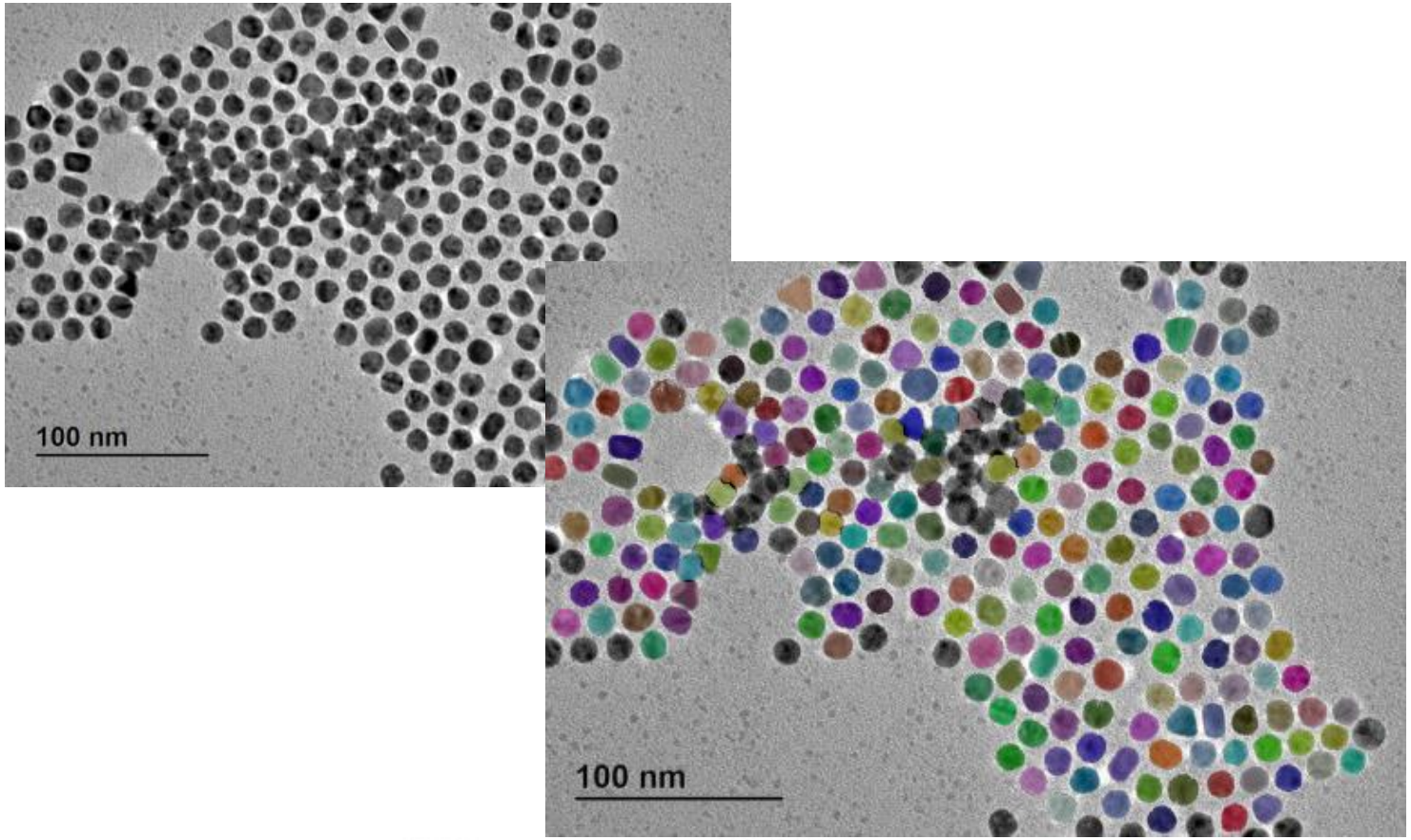
Development of a QNAR model [2]

Main steps:

1. Data collection and integration
2. Calculation of descriptors
3. Preprocessing and variable selection
4. Development of the *in silico* model for the prediction of the ENMs' biological effects
5. Model validation (internal, external) for testing predictive power of the model
6. Domain of applicability definition

Image Analysis

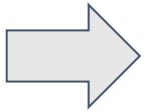
- Microscopy images
- Image processing
- Useful descriptors
  - Centroid X
  - Centroid Y
  - Circularity
  - Size
  - Eccentricity
  - Perimeter
  - Convexity etc.



Risk assessment platform [1]

Models open to the community: Development of a risk assessment web tools

QNAR models
Physicochemical descriptors
Image descriptors



Risk assessment platform
User-friendly
Ready-to-use
No need of previous programming knowledge
Ideal for experimentalists

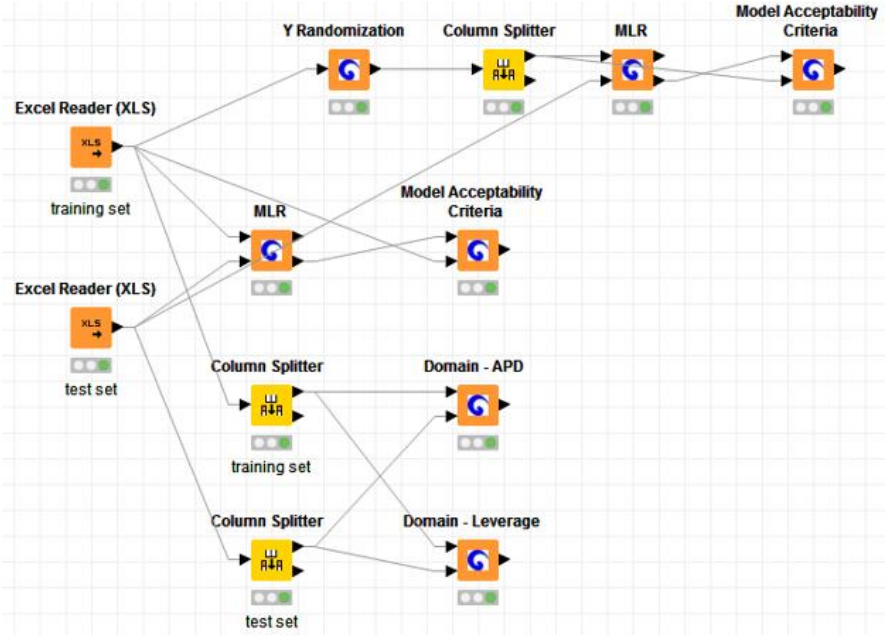
Risk assessment platform [2]

Enalos Nano/Cheminformatics Tools

Enalos+ nodes (through KNIME Analytics Platform)
Enalos Suite
Enalos Cloud Platform

KNIME Analytics Platform [1]

- A user-friendly and open-source platform that combines various software tools for data integration, processing, analysis, and exploitation
- Creation of a network of nodes
  - interact easily with the workflow
  - experiment with different methodologies in short-time
  - compare the results
  - have the complete supervision of the analysis process



KNIME Analytics Platform [2]

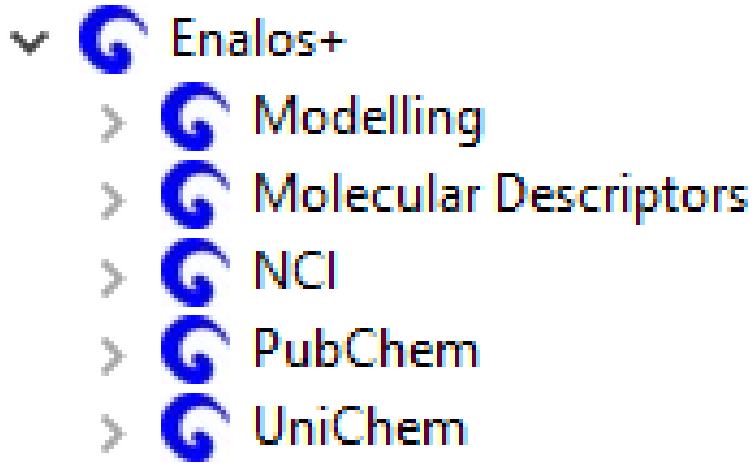
Enalos+ KNIME nodes [1]

- NovaMechanics Ltd made some very useful operations available as extensions for KNIME platform
- Enalos + nodes are fully compatible with other KNIME nodes
- Enalos+ nodes can be combined with custom made workflows and real time molecular descriptor calculations combined with state of the art modeling techniques (WEKA, R etc.)

http://enalosplus.novamechanics.com/

Enalos+ KNIME nodes [2]

- Data handling and preprocessing
- Calculation of molecular descriptors
- Modelling
- Testing the accuracy of the predictions
- Direct access to CIR (Chemical Identifier Resolver) through KNIME
- Direct access to the PubChem and UniChem databases and information acquisition for thousands of compounds



Molecular descriptors

- With molecular descriptors the chemical information contained in the molecule can be treated mathematically and can be used for modelling
- The structural characteristics can be directly linked with the biological or physicochemical properties of chemical compounds
- Mold2 (National Center for Toxicological Research of FDA), ideal for the calculation of molecular descriptors (777), encoding two-dimensional chemical structure information

Modelling nodes [1]

- Pre-processing nodes
  - Perform some simple but crucial procedures for handling the data and prepare them for modelling
  - Time-consuming procedures can be automated, eliminating significantly the effort and the time dedicated to them
  - Create New Molecules
  - Int 2 Double
  - Remove Column
  - Remove Duplicates

Modelling nodes [2]

- Partitioning nodes
  - Large datasets are difficult to handle and may cause computational problems
  - Reduction of the amount of data by dividing the initial dataset in smaller, representative subsets
  - Need of two representative subsets during an external model validation process (training and test sets)
  - Kennard and Stones, Sphere exclusion algorithms
- MLR node
  - Perform multiple linear regression to model the linear relationship between a dependent variable (target) and one or more independent variables (predictors)

Modelling nodes [3]

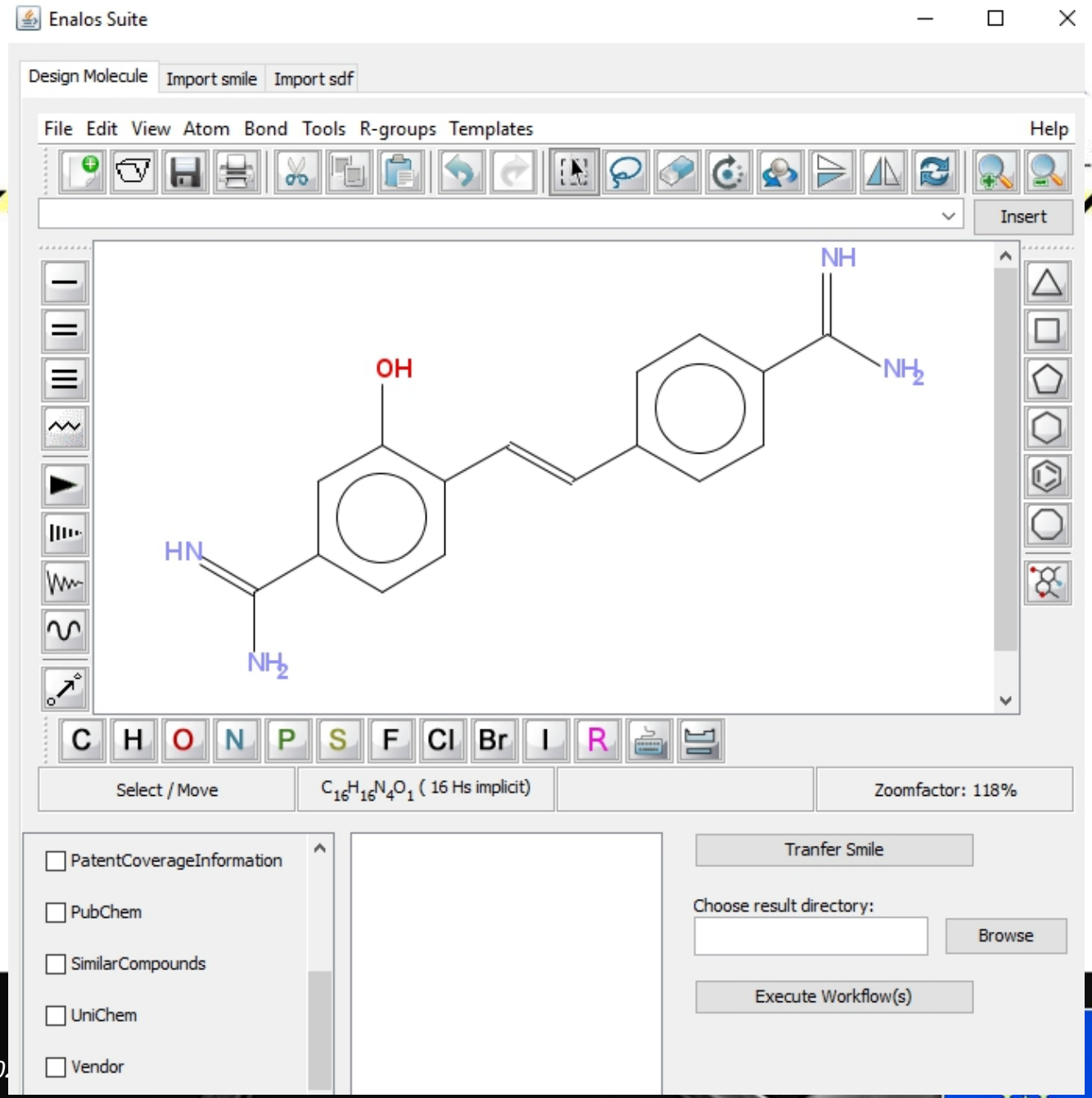
- Validation nodes
  - Techniques for the evaluation of the modelling
  - Define whether the generated predictions are reliable or not
  - Model Acceptability Criteria
  - Y-Randomization
- Domain nodes
  - Determination of the limits of the domain of applicability of the model
  - Predictions for only those compounds that fall into this domain may be considered reliable
  - Domain-Leverage, Domain-APD

Databases nodes

- The data analysis and modelling may be rather complicated
- Data from large collections and time may be lost while dealing with compatibility problems
- Database Enalos+ nodes give direct access to NCI, PubChem and UniChem chemical databases
- Importing data from databases via KNIME, offers a great flexibility
  - Direct analysis and handling with KNIME nodes
  - Fast and automated modelling

Enalos Suite

- Stand alone software that can package any predictive model developed by NovaMechanics in a completely custom made, independent platform
- The user can upload his own workflows and work via a user-friendly environment



Enalos Cloud Platform [1]

- An online, freely available toxicity and drug discovery platform
- Predictive models released as web services based on reliable, open source (KNIME, WEKA) and in-house developed software
- Address the need for reducing the amount of time and cost spent in experimental testing
- *In silico* methods and tools that produce accurate predictions for drug discovery and risk assessment of small molecules and novel ENMs

Enalos Cloud Platform [2]

http://www.insilicotox.com/

- Combined TNF-a & Solubility Prediction
- Aqueous Solubility Model
- A Risk Assessment Tool for the Virtual Screening of Metal Oxide Nanoparticles
- Prediction of MNPs Uptake in PaCa2 Cancer Cells

Enalos Cloud Nanoinformatics tools

Modeling of MNPs Uptake in $PaCa_2$ Cancer Cells

Virtual Screening of Metal Oxide Nanoparticles

Nanoparticles HepaRG classification

Modeling of MNPs Uptake in PaCa2 Cancer Cells [1]

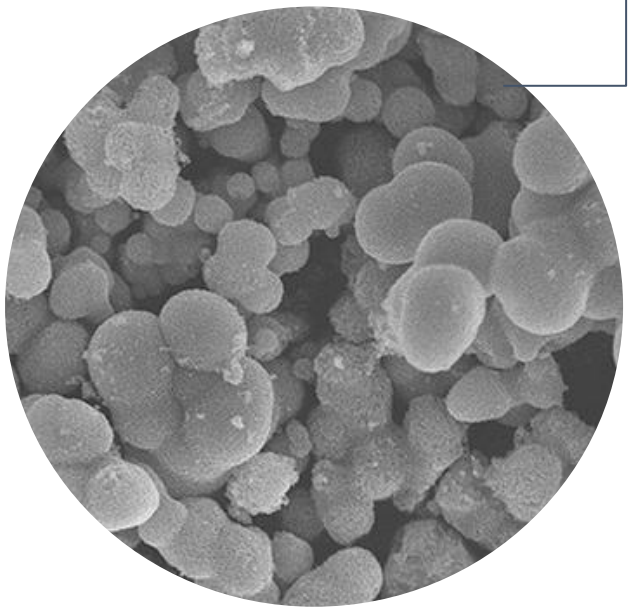
- Online toxicity predictions for coated iron oxide manufactured nanoparticles
- Prediction of the cellular uptake of NPs in pancreatic cancer cells
- Model development
  - Data available for 109 MNPs that have been synthesized and tested by the same group
  - Same NP core with different surface modifiers

Modeling of MNPs Uptake in PaCa2 Cancer Cells [2]

Same core

Different surface coatings

(Inputs)



Web service (model)

Predicted uptake (value)

Reliability of the prediction

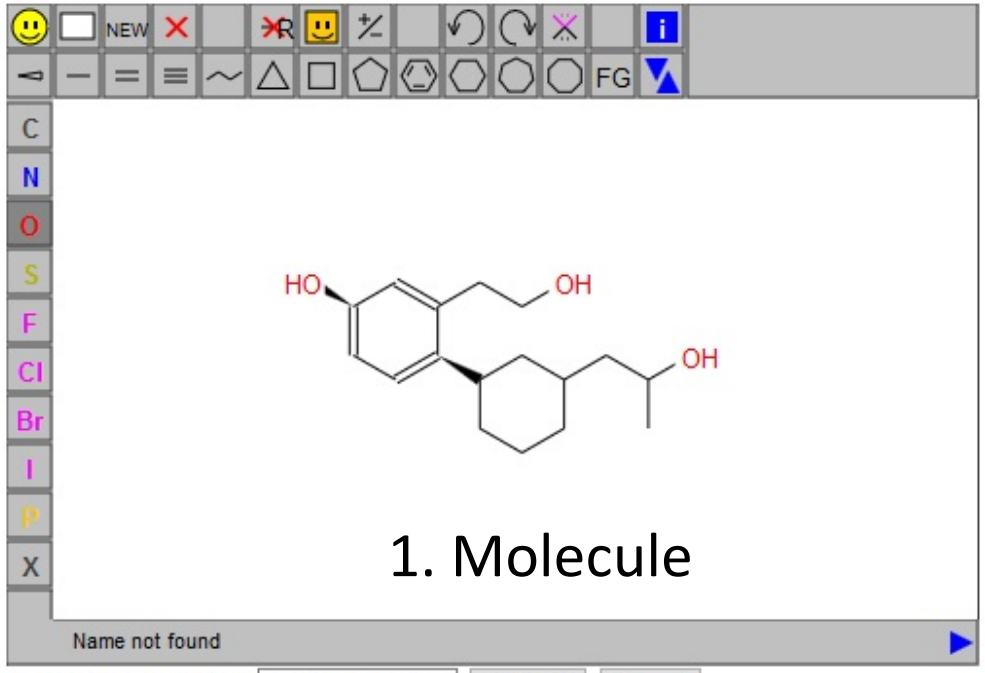
(Outputs)

http://enalos.insilicotox.com/QNAR_PaCa2/

Modeling of MNPs Uptake in PaCa2 Cancer Cells [3]

Prediction of MNPs Uptake in PaCa2 Cancer Cells

Design a molecule



1. Molecule

Select a workflow QNAR_PaCa2 Submit Reset

Enter SMILES separated by newlines

```
OC2CCCc1cccc12
C=C15CC(O)c2csccc12
C1CC2CCC3CCCC4CCC(C1)C2C34
C%10CCC(C19CCCC8CCC6CC5CCCC4CC2CCCC1C2C(C3CCCCC3)
(C45)C6(C7CCCCC7)C89)CC%10
c1cc2CCc3cccc4ccc(c1)c2c34
```

2. List of SMILES notations

Select a workflow QNAR_PaCa2 Submit Reset

Import an SDF file (.sdf)

Browse... No file selected. 3. SDF file

Select a workflow QNAR_PaCa2 Submit Reset

Modeling of MNPs Uptake in PaCa2 Cancer Cells [4]

Predicted values

Prediction of MNPs Uptake in PaCa2 Cancer Cells

Knime report *powered by Birt*

"PaCa2 cellular uptake (log10 [nanoparticles]/cell pM)"	"Domain of Applicability Prediction"
3.009	reliable
3.857	reliable
3.752	reliable
3.737	reliable

Date: Sep 20, 2017 4:13 PM Author: NovaMechanics Ltd 1 of 1

www.knime.org

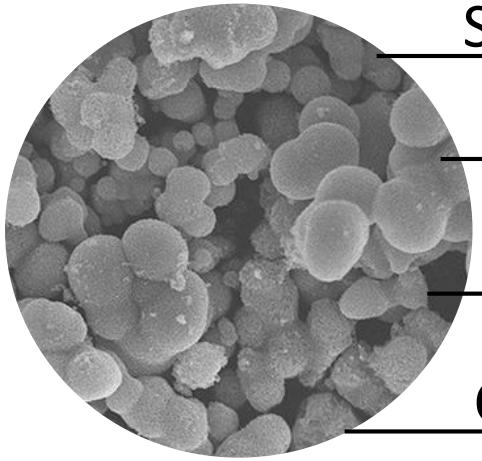
Reliability

G. Melagraki and A. Afantitis, "Enalos InSilicoNano platform: an online decision support tool for the design and virtual screening of nanoparticles", RSC Advances, vol. 4, pp. 50713-25, 2014

Virtual Screening of Metal Oxide NPs [1]

- Online toxicity predictions for Iron Oxide NPs
- Toxicity predictions (active/inactive) based on a set of indicated properties
- Model development
  - 44 iron oxide NPs (core: Fe_2O_3 and Fe_3O_4)
  - Coating: cross-linked dextran, polyvinyl alcohol, amphiphilic polymers
  - Sizes: 20-40, 74 nm
  - Relaxivities R1, R2 (mM-1s-1)
  - Zeta Potential (mV)
  - Evaluated in four cell types, four different assays
  - NPs classified: bioactive or inactive “safe” (threshold number of hits >= 4)

Virtual Screening of Metal Oxide NPs [2]



Size

Zeta Potential

Relaxivities R1, R2

Coating

Cross-linked dextran

Poly(vinyl alcohol) (PVA)

Other

Web service

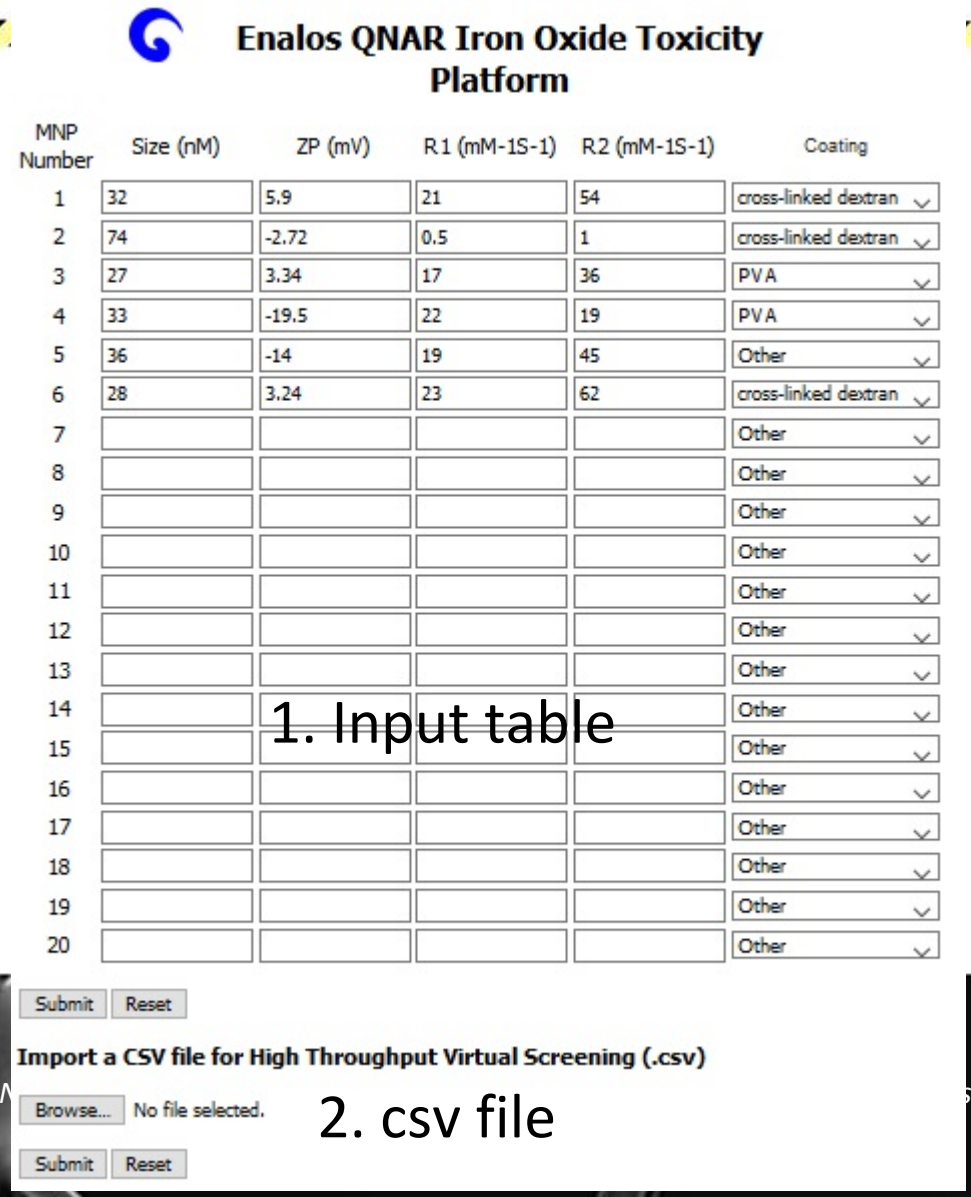
Predicted class active/inactive

Reliability of the prediction

(Outputs)

http://enalos.insilicotox.com/QNAR_IronOxide_Toxicity/

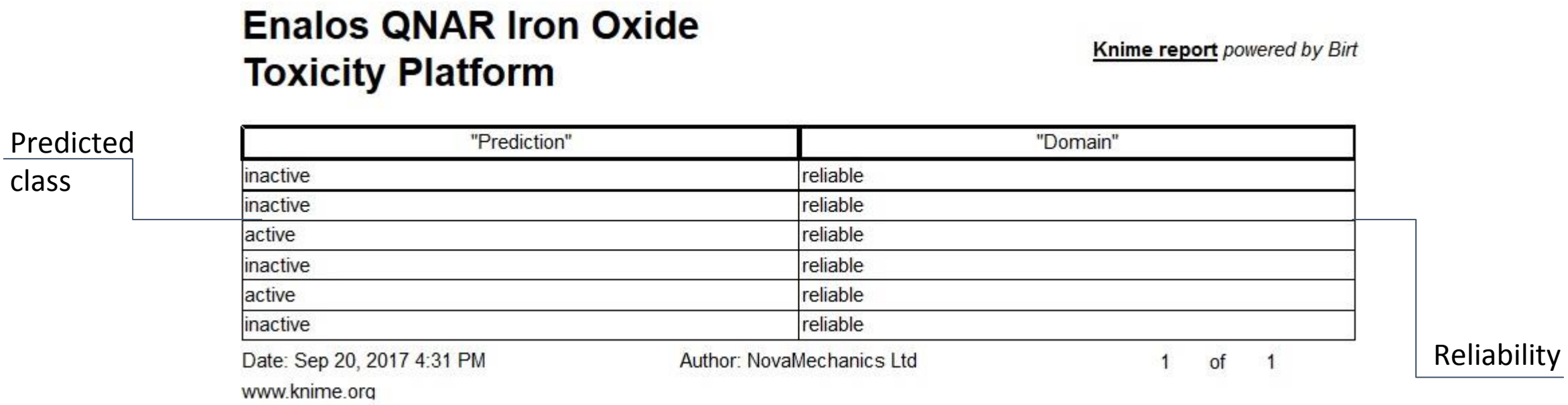
Virtual Screening of Metal Oxide NPs [3]



1. Input table

2. csv file

Virtual Screening of Metal Oxide NPs [4]



Enalos QNAR Iron Oxide Toxicity Platform

Knime report *powered by Birt*

Predicted class

"Prediction"	"Domain"
inactive	reliable
inactive	reliable
active	reliable
inactive	reliable
active	reliable
inactive	reliable

Reliability

Date: Sep 20, 2017 4:31 PM Author: NovaMechanics Ltd 1 of 1

www.knime.org

G. Melagraki and A. Afantitis, "A Risk Assessment Tool for the Virtual Screening of Metal Oxide Nanoparticles through Enalos InSilicoNano Platform", Current Topics in Medicinal Chemistry, vol. 15, no. 18, pp. 1827-36, 2015.

NPs HepaRG classification [1]

NANOGENTOOLS

Enalos Platform for Nanoparticles HepaRG Classification

MNP Number	Particle Core	Type of Coating	Particle non. Size (nm)	DLS (nm)	PDI	Zeta Potential (mV)	Electrophoretic mobiliy (μmcm/Vs)	Diameter (nm)	Shape
1	Ag	uncoated	15	27.36	0.471	-5.52	-0.3994	21.0966	Spherical
2	CeZrO2	uncoated	20	330.4	0.385	52.7	3.816	4.4	Spherical
3	CeO2	neutral	5	85.8	0.247	0.024	-0.7265	4.5	Spherical
4	ZnO	uncoated	5	4.229	0.362	0.04	0.02192	5.778	Spherical
5	Au	neutral	12	21.61	0.03	-16.4	-1.282	14.5	Spherical
6	SiO2	anionic	60	65.22	0.03	-12.9	-1.025	56.8	Spherical
7	TiO2	neutral	10	3109	0.177	17.1	1.344	24.07551	Faceted
8	TiO2	uncoated	20	403	0.81	25.3	1.981	37.44436	Nanorods Various shapes
9	CeO2	neutral	30	93.7	0.152	5.7	0.7289	30.8	Faceted
10	CuO	neutral	20	8.119	0.406	-0.59	-0.04256	12.144	Spherical
11	ZnO	uncoated							Spherical
12	ZnO	uncoated							Spherical
13	ZnO	uncoated							Spherical
14	ZnO	uncoated							Spherical
15	ZnO	uncoated							Spherical
16	ZnO	uncoated							Spherical
17	ZnO	uncoated							Spherical
18	ZnO	uncoated							Spherical
19	ZnO	uncoated							Spherical
20	ZnO	uncoated							Spherical

Submit Reset

- Particle core
- Type of coating
- Particle size
- DLS
- PDI
- Zeta potential
- Electrophoretic mobility
- Diameter
- Shape

NPs HepaRG classification [2]

Prediction for 4 different endpoints

http://enalos.insilicotox.com/HepaRG/

Predicted HepaRG NP classification

Knime report *powered by Birt*

"Prediction (ClassCC)"	"Prediction (ClassVCC)"	"Prediction (ClassCMD)"	"Prediction (ClassMMP)"	"Prediction"
Toxic	Toxic	Toxic	Non toxic	reliable
Non toxic	Non toxic	Non toxic	Non toxic	reliable
Non toxic	Non toxic	Non toxic	Non toxic	reliable
Toxic	Toxic	Toxic	Non toxic	reliable
Non toxic	Non toxic	Non toxic	Non toxic	reliable
Non toxic	Non toxic	Non toxic	Non toxic	reliable
Toxic	Toxic	Toxic	Toxic	reliable
Toxic	Toxic	Toxic	Non toxic	reliable
Non toxic	Non toxic	Non toxic	Non toxic	reliable
Toxic	Toxic	Toxic	Toxic	reliable

Date: Feb 15, 2017 10:45 AM Author: NanoMILE - NovaMechanics Ltd 1 of 1

www.knime.org

NPs HepaRG classification [3]

Out of the model's domain of applicability!

Enalos Platform for Nanoparticles HepaRG Classification

MNP Number	Particle Core	Type of Coating	Particle non. Size (nm)	DLS (nm)	PDI	Zeta Potential (mV)	Electrophoretic mobiliy (μmcm/Vs)	Diameter (nm)	Shape
1	Ag	uncoated	15	27.36	0.471	-5.52	-0.3994	21.0966	Spherical
2	CeZrO2	uncoated	20	330.4	0.385	52.7	3.816	4.4	Spherical
3	CeO2	neutral	5	85.8	0.247	0.024	-0.7265	4.5	Spherical
4	ZnO	uncoated	5	4.229	0.362	0.04	0.02192	5.778	Spherical
5	Au	neutral	12	21.61	0.03	-16.4	-1.282	14.5	Spherical
6	SiO2	anionic	60	65.22	0.03	-12.9	-1.025	56.8	Spherical
7	TiO2	neutral	10	3109	0.177	17.1	1.344	24.07551	Faceted
8	TiO2	uncoated	20	403	0.81	25.3	1.981	37.44436	Nanorods Various shapes
9	CeO2	neutral	30	93.7	0.152	5.7	0.7289	30.8	Faceted
10	CuO	neutral	200	8.119	0.406	-0.59	-0.04256	12.144	Spherical
11	ZnO	uncoated							Spherical
12	ZnO	uncoated							Spherical
13	ZnO	uncoated							Spherical
14	ZnO	uncoated							Spherical
15	ZnO	uncoated							Spherical
16	ZnO	uncoated							Spherical
17	ZnO	uncoated							Spherical
18	ZnO	uncoated							Spherical
19	ZnO	uncoated							Spherical
20	ZnO	uncoated							Spherical

Submit Reset

NPs HepaRG classification [4]

Predicted HepaRG NP classification

Knime report *powered by Birt*

"Prediction (ClassCC)"	"Prediction (ClassVCC)"	"Prediction (ClassCMD)"	"Prediction (ClassMMP)"	"Prediction"
Toxic	Toxic	Toxic	Non toxic	reliable
Non toxic	Non toxic	Non toxic	Non toxic	reliable
Non toxic	Non toxic	Non toxic	Non toxic	reliable
Toxic	Toxic	Toxic	Non toxic	reliable
Non toxic	Non toxic	Non toxic	Non toxic	reliable
Non toxic	Non toxic	Non toxic	Non toxic	reliable
Toxic	Toxic	Toxic	Toxic	reliable
Toxic	Toxic	Toxic	Non toxic	reliable
Non toxic	Non toxic	Non toxic	Non toxic	reliable
Toxic	Toxic	Toxic	Toxic	unreliable

Date: Feb 15, 2017 11:06 AM Author: NanoMILE - NovaMechanics Ltd 1 of 1

www.knime.org

NovaMechanics Ltd

www.novamechanics.com

info@novamechanics.com

Enalos+ Tools

http://enalosplus.novamechanics.com

Thank you!

The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 691095.