

Data strategy and relevance for Nanoinformatics and Nanosafety projects

Vladimir Lobaskin
University College Dublin

Project legacy

Concepts

- Mechanism-based toxicity assessment paradigm



Methods and SOPs

- NM labelling, tracking, post-uptake characterisation
- ALI exposure imitation, aerosolisation, exposure quantification
- Bionano interaction modelling
- Pathway analysis, GRN analysis, omics data processing



Models

- QSARs for MIE/KE
- Protein adsorption and corona models



Data

- In vivo: histology, omics, AOPs, GRNs, AOPs, MIEs, KEs
- In vitro: cell culture tox assays, NM tracking, NM corona, ROS, surface tension
- In silico: NM properties, bionano interactions



SmartNanoTox outcomes: Toxicology

- Description of 5 respiratory AOPs, identified KE/MIE, tools for KE/MIE – **AOP Wiki, NanoCommons**
- Gene expression profiles for *in vivo* respiratory exposure – **GEO database**
- Novel ALI systems to imitate realistic exposure conditions - **Vitrocell**
- Novel analysis protocols for inference of GRNs from transcriptomics data, identification of Core Regulatory Genes – **NanoCommons**



Available via project website www.smartnanotox.eu

Project outcomes: Nanoparticles

- Novel NM labelling techniques – **JSI, SNT website**
- Novel protocols for corona analysis – **SNT website**
- Novel algorithms for image analysis / colocalization – **JSI**
- Protein corona-based NM fingerprints – **NanoCommons KB**
- NM tracking techniques, post-uptake characterisation data - **JSI**
- Atomistic, coarse-grained force fields for common materials (30 materials) – **NanoSolveIT, OpenKIM, SNT website**
- Multiscale simulation tools for bionano-interface – **NanoCommons KB**
- Novel advanced NM and protein descriptors (over 30 new descriptors, 60 materials) – **NanoCommons KB**
- Publicly available database of bionano interactions – **NanoCommons KB**

Project outcomes: Industry/Regulation

- Novel toxicity endpoints bound to in vivo AOPs
- Novel in vitro assays targeting MIE/KE based on reporter gene
- Elucidation of toxicity mechanisms for oxides, carbonaceous materials
- Creation of basis for read-across and safety by design through identification of NM properties of concern
- Mechanism-aware QSARs relating NM properties to biological activity – **SNT Website**



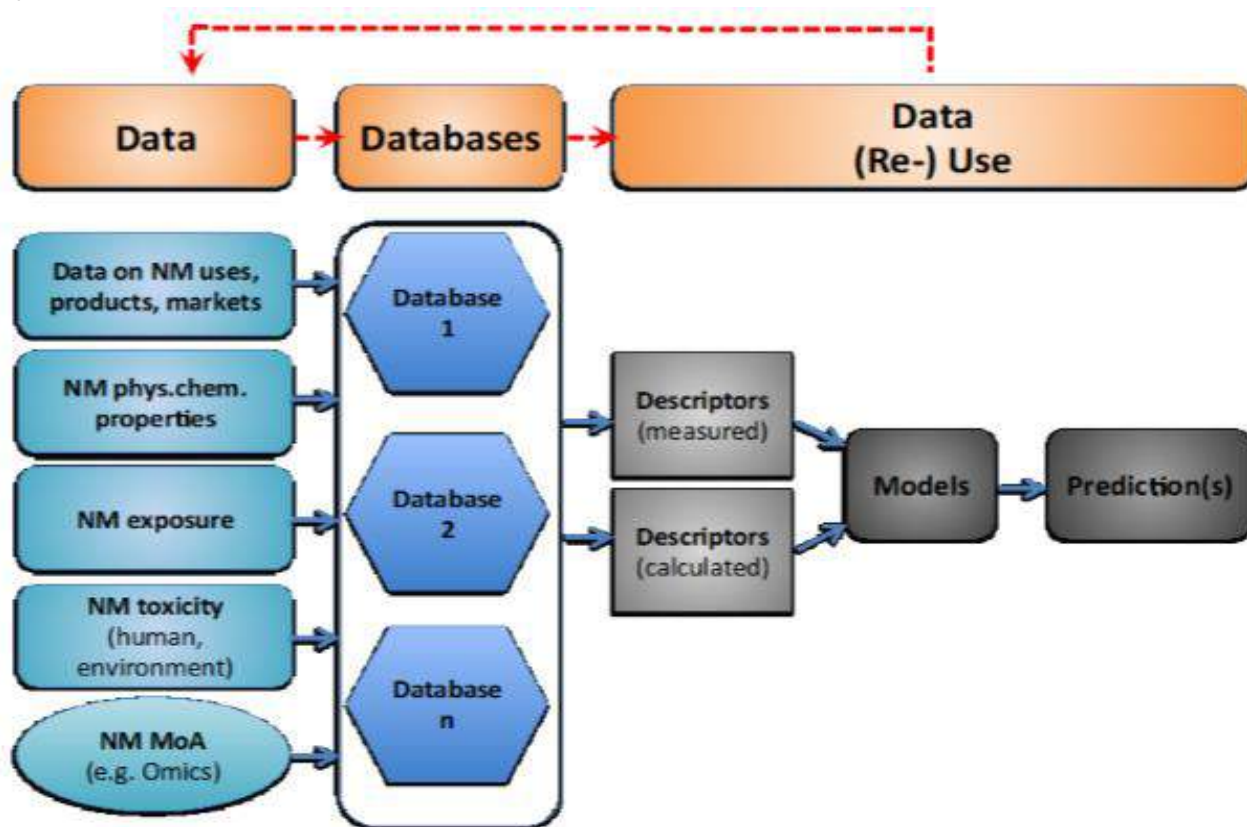
SmartNanoTox materials

Smart Tools for Gauging Nano Hazards

- Titania (TiO_2): 16 NMs – rutile, anatase, spheres, tubes
- Quartz and Silica: 3 NMs
- Metal oxides: 7 NMs Fe_2O_3 , ZnO , $\text{Ni}_x\text{Fe}_y\text{Zn}_z\text{O}_u$
- Carbon black: 2 NMs
- Carbon nanotubes: 28 SWCNTs and MWCNTs
- Graphene: 3 NMs
- Asbestos: Crocidolite

Nanoinformatics

Relevant NM characterisation data are necessary to build predictive models



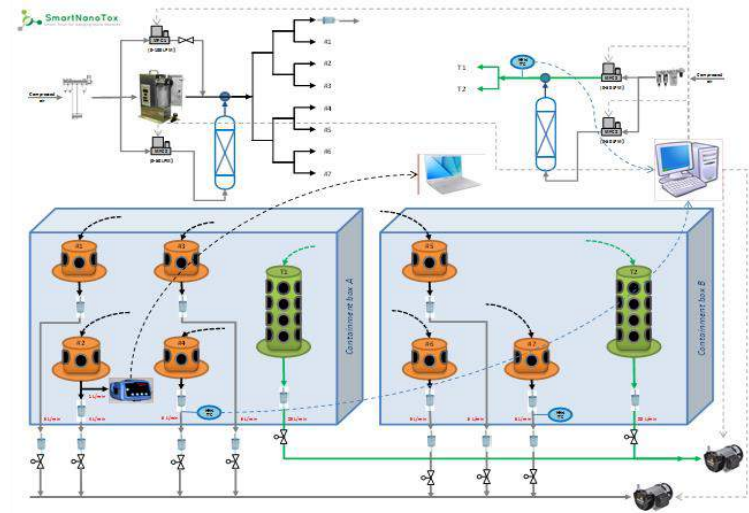
Contribution to nanoinformatics

- NM properties measured/calculated for many NMs
- Novel NM descriptors: bionano interface, surface activity
- Methods of in silico NM characterisation: ab initio multiscale models
- Mechanism-aware QSARs relating NM properties to biological activity MIE/KE
- Grouping of materials based on their bio-relevant activities

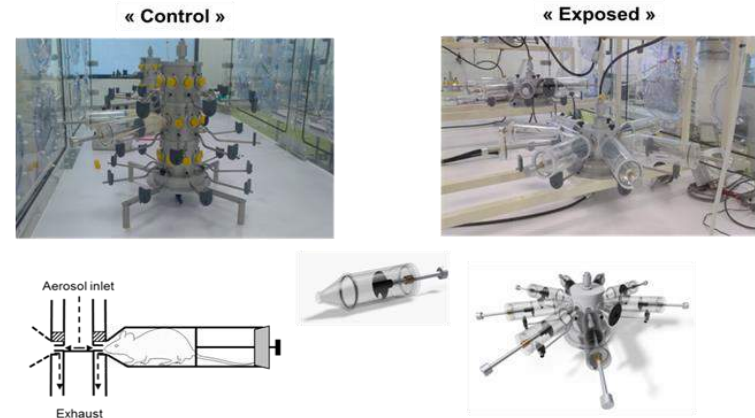
SmartNanoTox Data: in vivo

- Inhalation: rats, mice
- Instillation: mice
- 3 time points
- 3 concentrations
- 15+ NMs

A



B

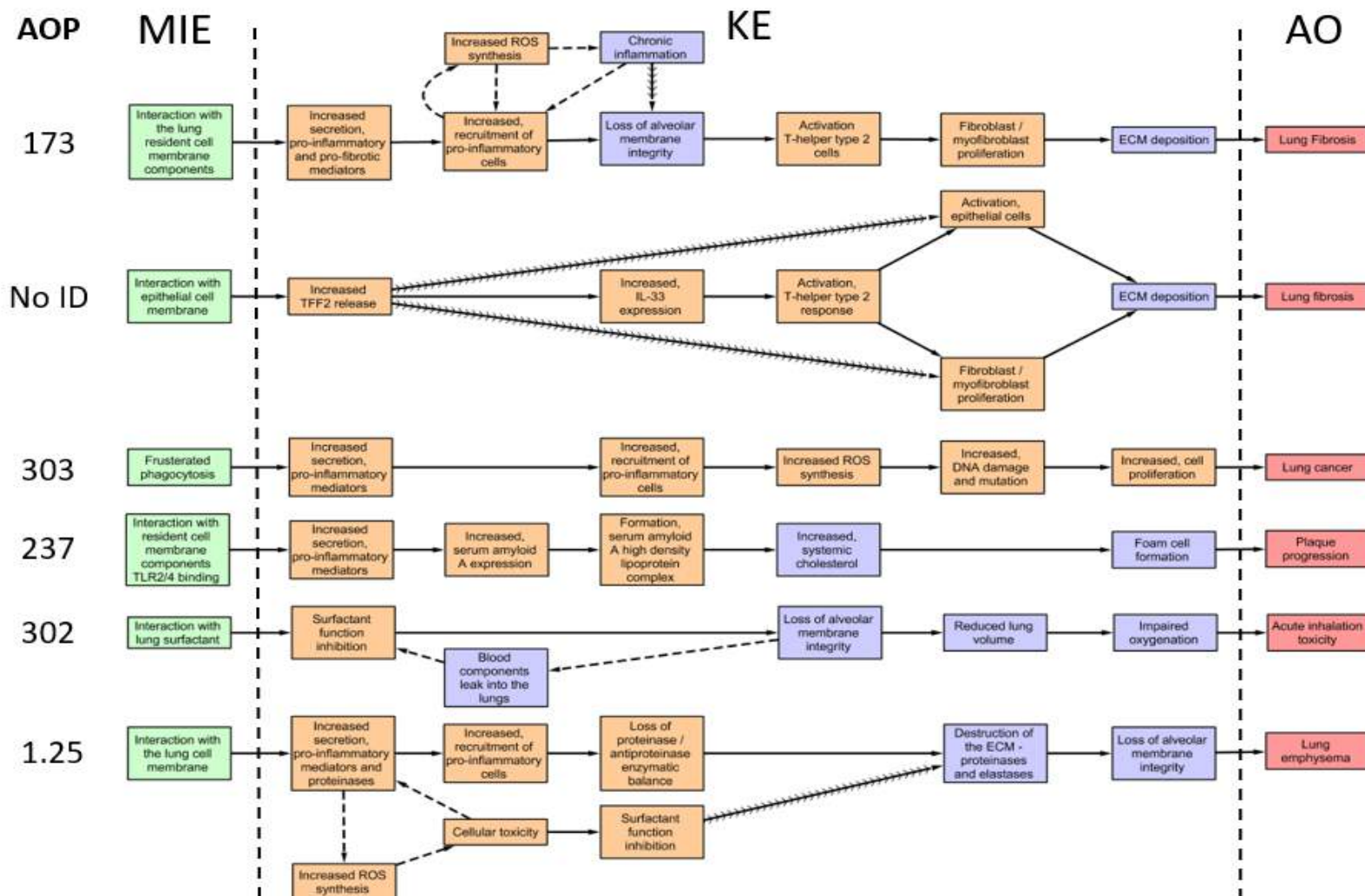


SmartNanoTox Data: in vivo

Study	Activity	Activity Description	Type of data for sharing	Shared Format	Archive
NMs / Tissues	'Omics	Studies for in vivo reaction to inhalation/ instillation of NMs: gene expression, metabolomics, proteomics	Candidate Gene and Protein IDs	MaxQuant Data CSV*	PRIDE, GEO, Zenodo
Inhalation / IT instillation of NMs	In vivo tox tests in mice	Inhalation experiments: defined NMs aerosolisation, & investigation of interactions of NM within the lung lining fluid, alveolar epithelial cells & macrophages	Toxicological endpoints - Tissue for histological evaluation and gene profiling, analysis of biological endpoints	data sheets, CSV*	SNT, Zenodo
Lung-specific or respiratory AOP	Toxicity Pathway Identification	Identify pathways for lung diseases from GEO & biological processes perturbed by NM uptake	Candidate Pathways	CSV*	GEO, Zenodo
IT instillation of NMs	In vivo tox tests in mice	Instillation experiments: Investigation of interactions of NM within the lung lining fluid, alveolar epithelial cells and macrophages	Toxicological endpoints/Tissue for histological evaluation & gene profiling, analysis of biological endpoints	CSV*	Zenodo
Lung tissue	Histology evaluation	Tissue analysis of in vivo samples	**Verbal** evaluation, Semi-quantitative scores	CSV*, images**, MDs	SNT, Zenodo

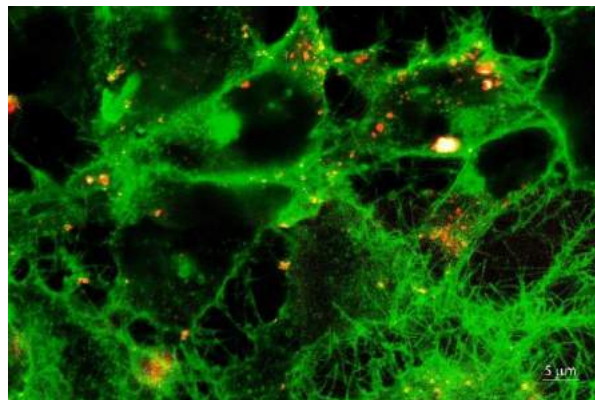
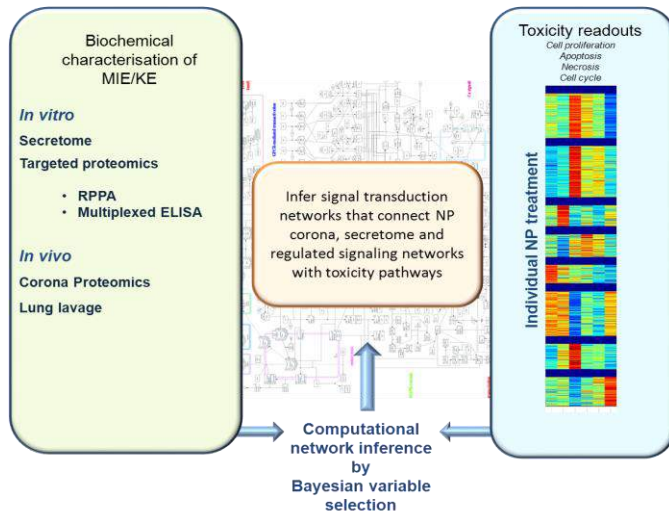
Network of NM-induced AOPs

Figure 1

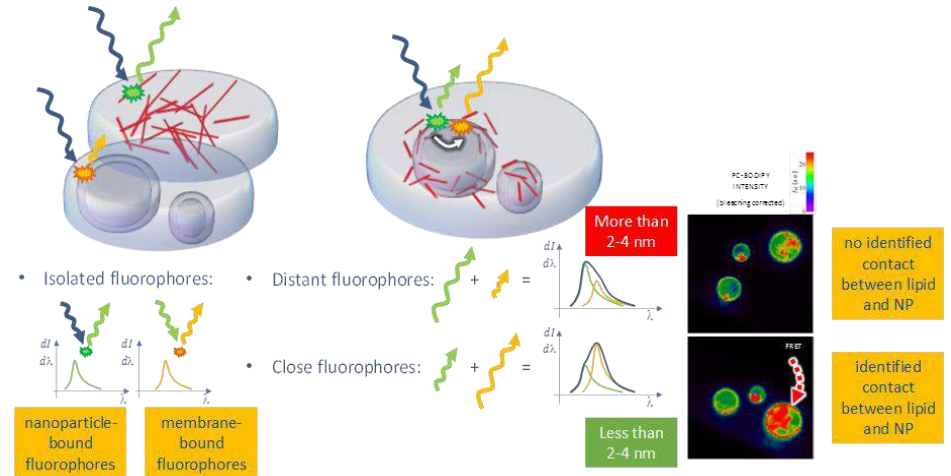


SmartNanoTox Data: in vitro

Analysis workflow



How nanoparticle-lipid/protein contact will be identified?



SmartNanoTox Data: in vitro

Study	Activity	Activity Description	Type of data for sharing	Shared Format	Archive
NMs	Lipidomics	Content & evolution of the NM biomolecular corona & lipid wrap/NM containing complexes from in vivo exp.	MS derived Lipid IDs	Spectral Data CSV*	SNT, Zenodo
Lung cells	In vitro tox tests in murine cells	Cell-line exposure experiments using Vitrocell exposure systems	Cytotoxicity and biological endpoints (gene-/protein-level)	data sheets, CSV*	SNT, Zenodo
NMs	Surface Tension Experiments	Investigation of NM interaction with the lung alveolar fluid and lung integrity	Surface tension profiles, dose-response curves Effect of NMs on lung surfactant surface tension, inhibitory dose	CSV*	Zenodo
Labelled NPs	FMS	NM modification for efficient tracking (fluorescent/isotopic labelling)	fluorescence spectra - procedures for labelling nanoparticles & characterize the efficiency of labelling	spectra, synthetic pathways, CSV*	LBF Website, Zenodo
Labelled NPs / cell lines	FMS, 3D STED spectral data	NM modification for efficient tracking (fluorescent/isotopic labelling)	2D & 3D stack of 2D images (4D & 5D sets of data) - images, evaluation of affinity constants, lifetimes of states, characteristic times of membrane changes	Images**, text data*, Video	LBF Website, Zenodo
NPs / BALF / Lipid dispersion	In vitro 'NP' to lipid affinity	Analysis of NM-lipid interaction for labelled NMs	autocorrelation curves analysis & stack of images - images, evaluation of affinity constants	Text data, CSV*, Images**	LBF Website, Zenodo
Lung cells	In vitro tox tests rat & human cells	Analysis of cell responses to NMs	Cytotoxicity and biological endpoints (gene-/protein-level), included transcriptomics	data sheets CSV*	SNT, Zenodo

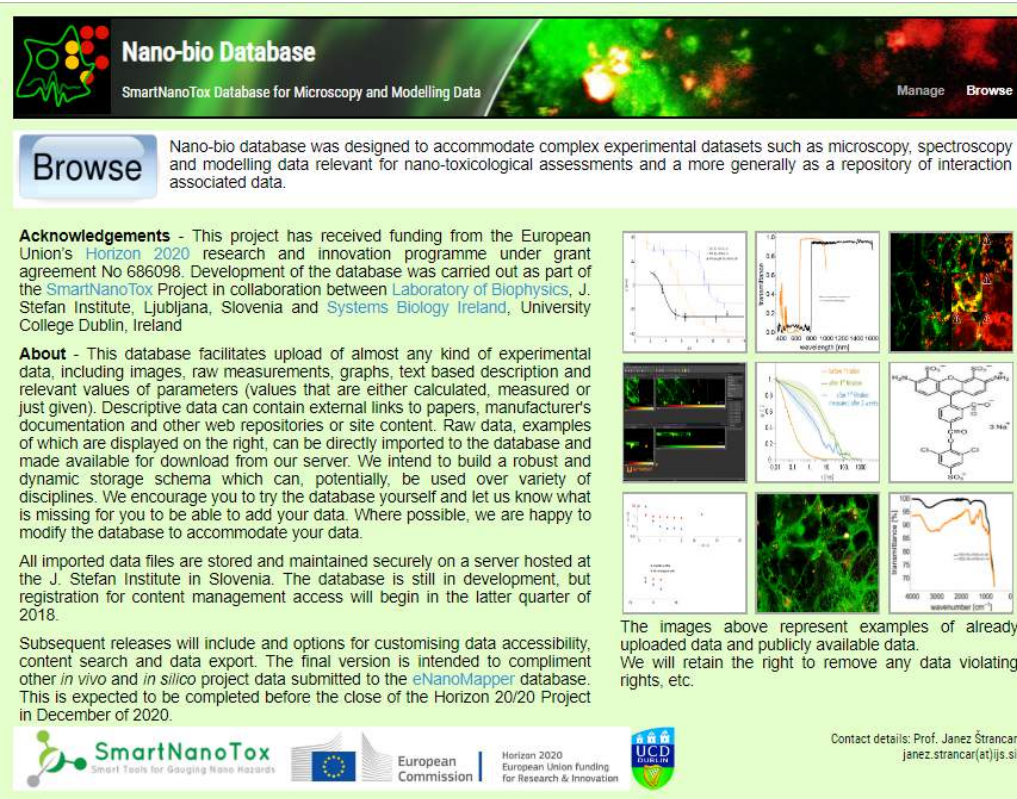
LBF XML Data Schema

Underlying Schema

- Comprehensive set of XML tags for describing location, image, text & numerical data.
- XML paired with styling info. should render in any browser if/when downloaded.

Data Service

- Current version (public) lbfnanobiodatabase.ijs.si.
- Components can be referred to explicitly or searched via URI.
- An unreleased version facilitates data search and personalised user 'libraries'.
- Libraries can be shared between users (service) or made public & | downloadable.



Nano-bio Database
SmartNanoTox Database for Microscopy and Modelling Data

Manage Browse

Browse Nano-bio database was designed to accommodate complex experimental datasets such as microscopy, spectroscopy and modelling data relevant for nano-toxicological assessments and a more generally as a repository of interaction associated data.

Acknowledgements - This project has received funding from the European Union's [Horizon 2020](#) research and innovation programme under grant agreement No 686098. Development of the database was carried out as part of the [SmartNanoTox](#) Project in collaboration between [Laboratory of Biophysics](#), J. Stefan Institute, Ljubljana, Slovenia and [Systems Biology Ireland](#), University College Dublin, Ireland

About - This database facilitates upload of almost any kind of experimental data, including images, raw measurements, graphs, text based description and relevant values of parameters (values that are either calculated, measured or just given). Descriptive data can contain external links to papers, manufacturer's documentation and other web repositories or site content. Raw data, examples of which are displayed on the right, can be directly imported to the database and made available for download from our server. We intend to build a robust and dynamic storage schema which can, potentially, be used over variety of disciplines. We encourage you to try the database yourself and let us know what is missing for you to be able to add your data. Where possible, we are happy to modify the database to accommodate your data.

All imported data files are stored and maintained securely on a server hosted at the J. Stefan Institute in Slovenia. The database is still in development, but registration for content management access will begin in the latter quarter of 2018.

Subsequent releases will include options for customising data accessibility, content search and data export. The final version is intended to compliment other *in vivo* and *in silico* project data submitted to the [eNanoMapper](#) database. This is expected to be completed before the close of the Horizon 2020 Project in December of 2020.

The images above represent examples of already uploaded data and publicly available data. We will retain the right to remove any data violating rights, etc.

Contact details: Prof. Janez Štrancar
janez.strancar@ijs.si

SmartNanoTox
Smart Tools for Gauging Nano Hazards

European Commission | Horizon 2020
European Union Funding
for Research & Innovation

UCD
University College Dublin



Biological (Pending)

1. Biological_system_v2_BKmod

Analysis

1. Analysis_type_v1

Filter Sets

1. Filter_set_v1

Labels

1. AlexaFlour

Lasers

1. Lasers_v3

Measurements

1. Measurement_v1_01

Media

1. Media_v1

Methods

1. Methods_v1_BK_08

2. Methods_v1_HM_03_BKmod

Observations

1. Observations_v1

References

1. References_v1

```
--<root>
- <biosystem_class id="[AUTO]" type="in vitro">
- <!--
  HM - other possibilities: in vivo, ex vivo, in silico etc. we only use in vit:
-->
- <biological_system id="[AUTO]" type="cell line">
- <!--
  HM - other possibilities: liposomes = model membranes, ...
-->
  <cell_line id="[AUTO]" type="LA-4"/>
- <source_material id="[AUTO]">
  *
  <!-- 20/08 - BK - we get cells from someone -->
  <industry_standard_name id="[AUTO]">LA-4</industry_standard_name>
- <!--
  all numbers should be suffix (?the small number below)
-->
  <synonym>LA-4 epithelial cells</synonym>
  <lab_identifier>LA-4 cells</lab_identifier>
- <!--
  in-house name used in lab books. If standardised: NO Methods / If made/mod:
-->
  <method/>
  <!-- Don't worry too much about this this week. -->
  <reference id="[AUTO]" type="" source="Helmholtz Institute - Munich"/>
  <raw_material supplier="Carola Endes/Tobias Stoerger"/>
</source_material>
<synonym>LA-4 epithelial cells</synonym>
*
<reference id="[AUTO]" type="website" source="" uri=""/>
<!-- 20/08 - BK - do we need this? -->
```



Biological (Pending)

1. [Biological_system_v2_BKmod](#)

Analysis

1. [Analysis_type_v1](#)

Filter Sets

1. [Filter_set_v1](#)

Labels

1. [AlexaFlour](#)

Lasers

1. [Lasers_v3](#)

Measurements

1. [Measurement_v1_01](#)

Media

1. [Media_v1](#)

Methods

- 1. [Methods_v1_BK_08](#)
- 2. [Methods_v1_HM_03_BKmod](#)

Observations

1. [Observations_v1](#)

References

1. [References_v1](#)

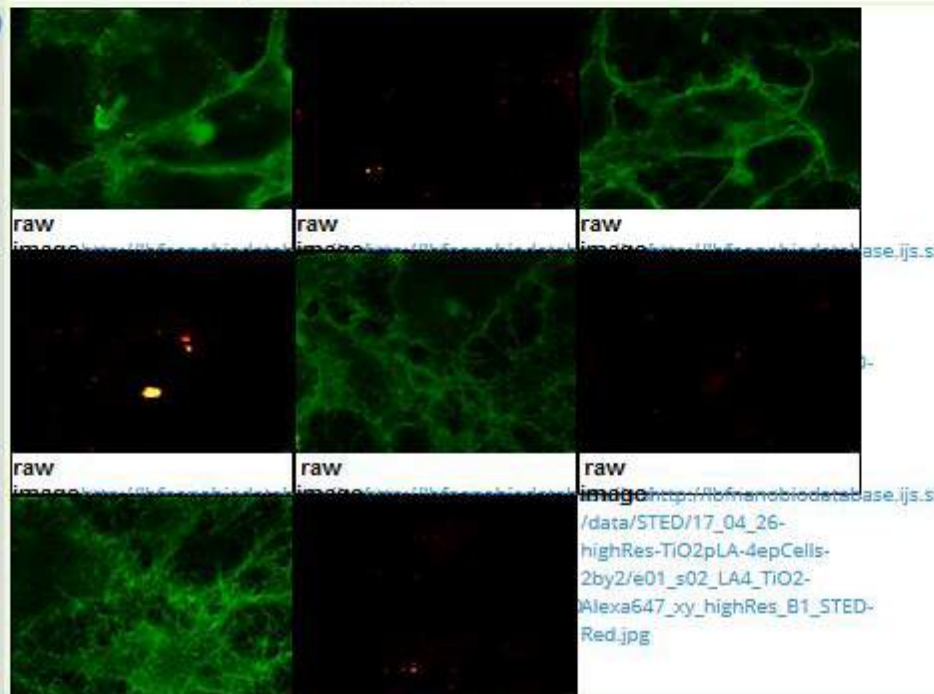
Method Class **Optical microscopy**

ID:[AUTO] Method **STED microscopy**

Observation(s)

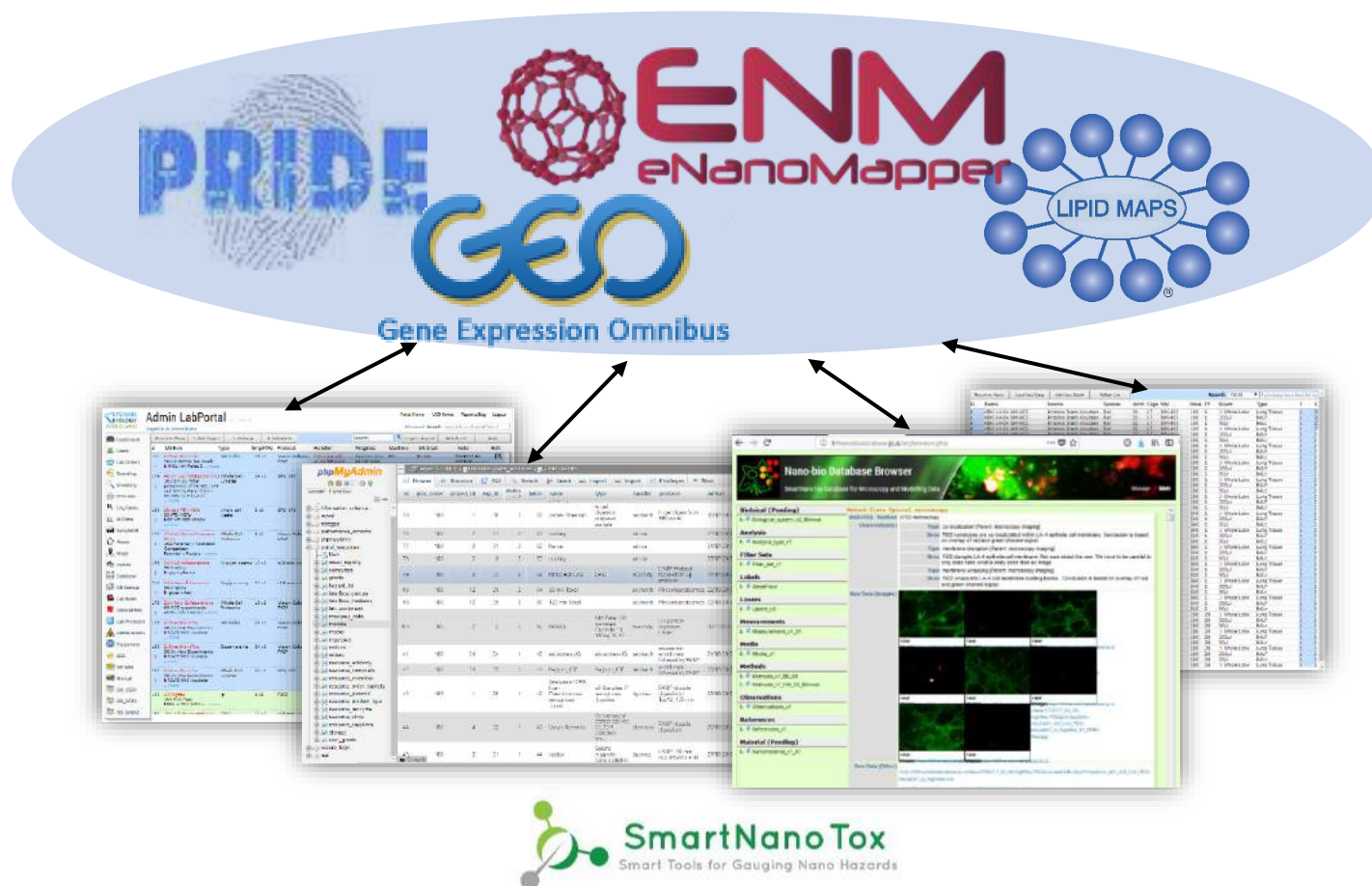
Type	Desc
co-localization [Parent: microscopy imaging]	TiO2 nanotubes are co-localized within LA-4 epithelia cell membrane. Conclusion is based on overlay of red and green channel signal.
membrane disruption [Parent: microscopy imaging]	TiO2 disrupts LA-4 epithelia cell membrane. Not sure about this one. We have to be careful to only state here what is really seen from an image
membrane wrapping [Parent: microscopy imaging]	TiO2 wraps into LA-4 cell membrane building blocks . Conclusion is based on overlay of red and green channel signal.

Raw Data (Images)



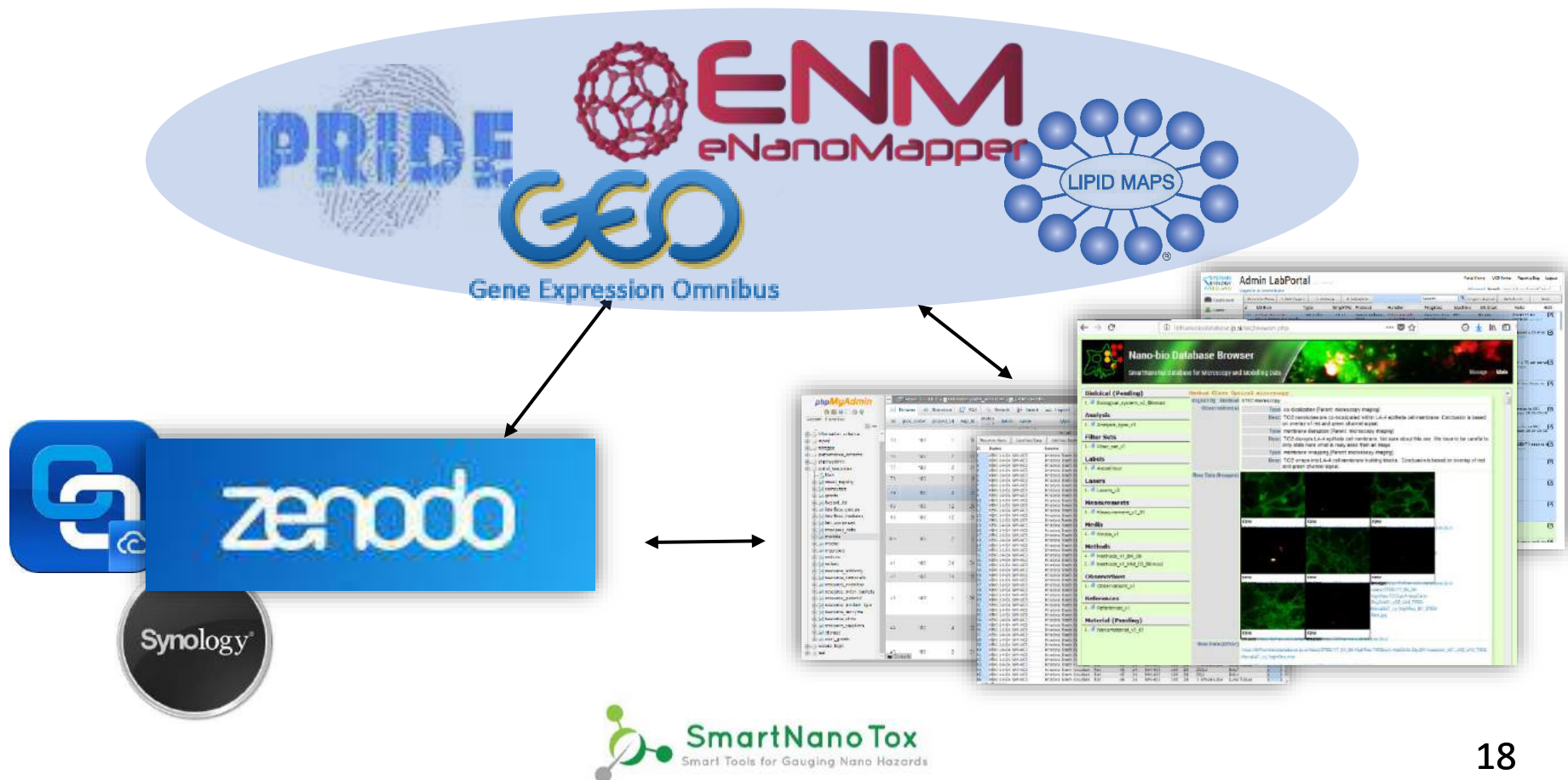
Development of Novel Data Services

- SNT data which is not eNM/NC KB compliant and lack an established or centralised data repository can be transformed into XML databases and referenced via embedded URIs.

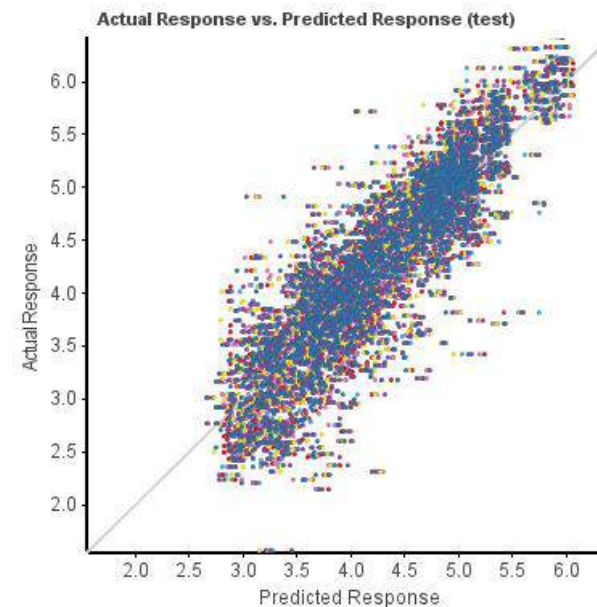
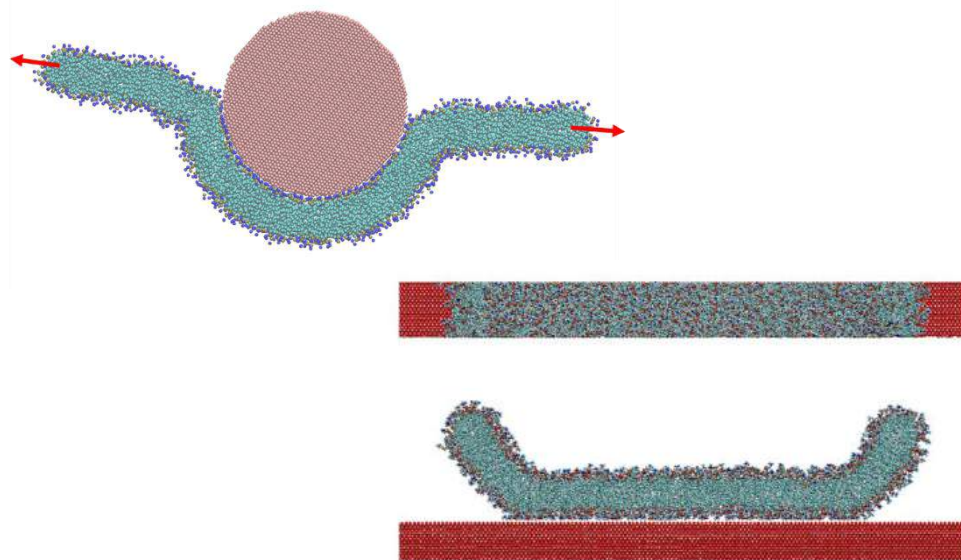
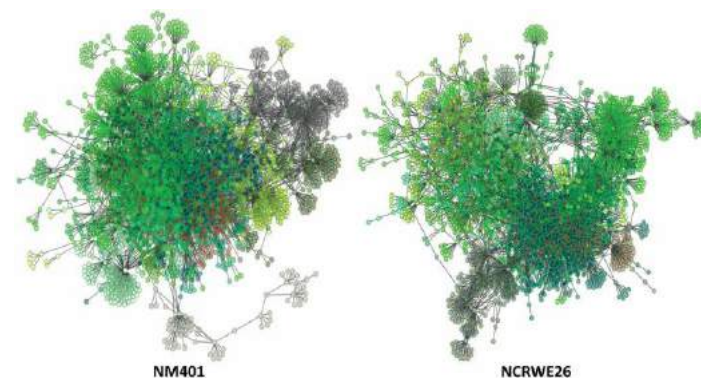
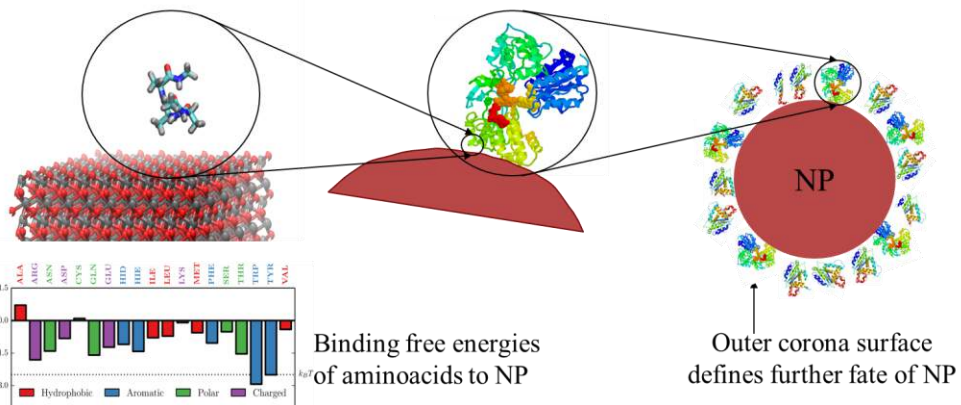


Supplementary File Storage

- 200+ GB Simulation data and 2+ TB Proteomics not all shared via research repositories.
- Large volume data storage will be hosted via NAS & Zenodo File storage services.



SmartNanoTox Data: in silico

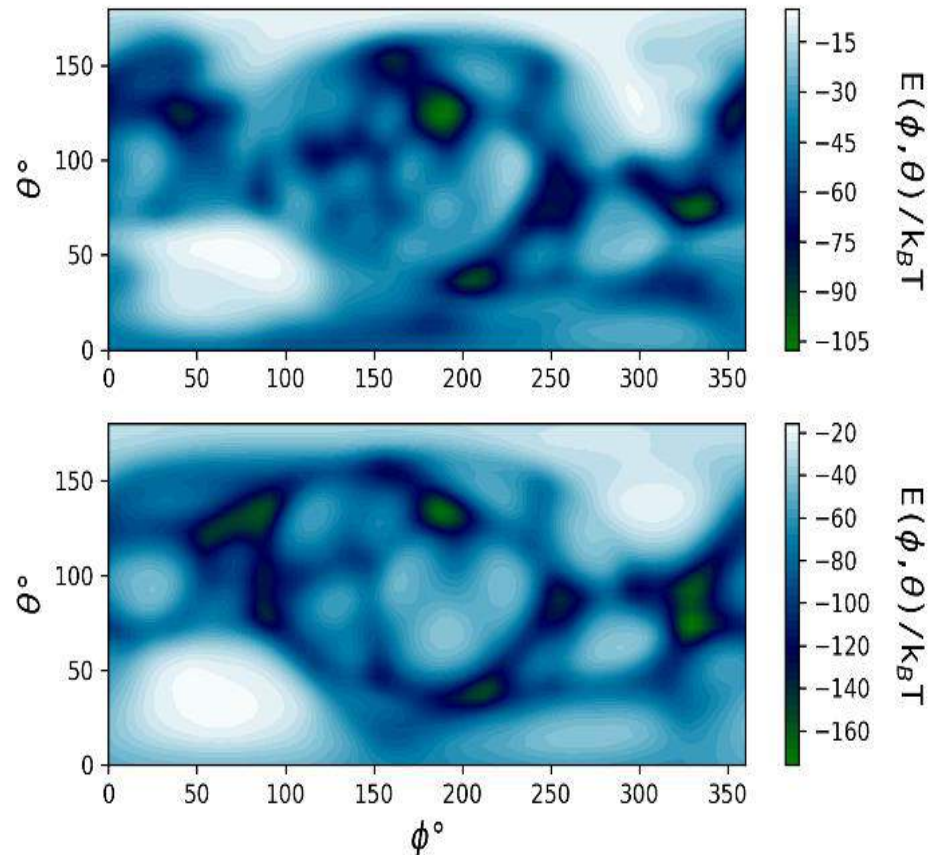
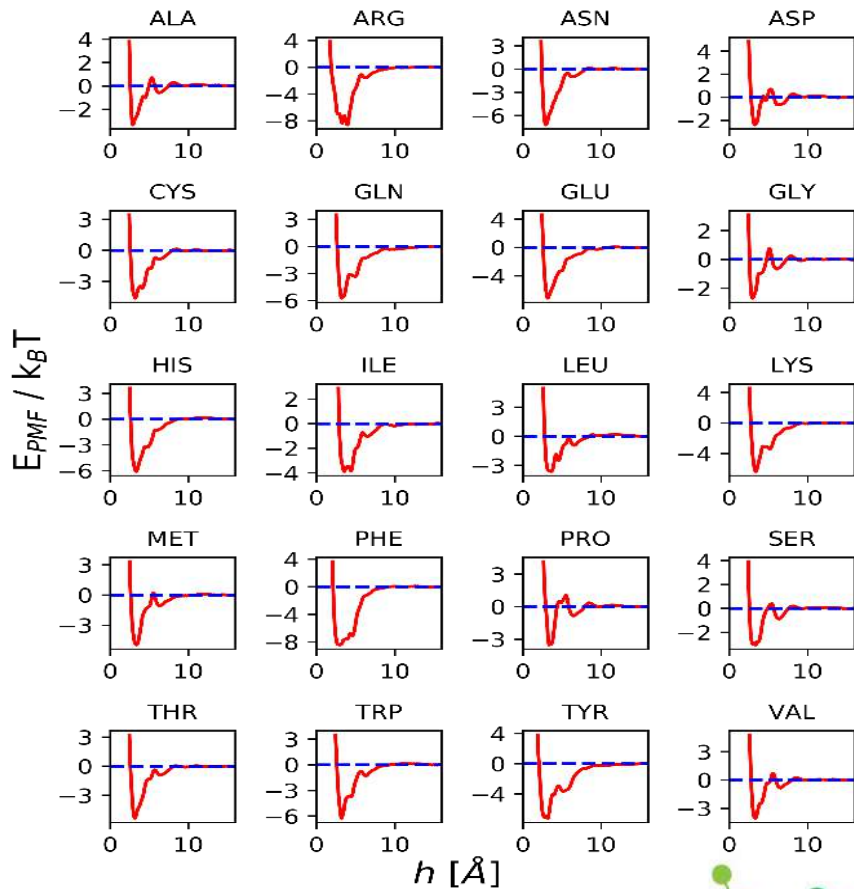


SmartNanoTox Data: in silico

Study	Activity	Activity Description	Type of data for sharing	Shared Format	Archive
Proteins	Molecular simulation	Prediction of 3D structure of proteins using i-Tasser software, calculation of molecule characteristics	coarse-grained models, sequence descriptors, structure descriptors, adsorption energies	CSV*	SNT, Zenodo
NMs	Molecular simulation	Quantum-chemical modelling of NM, calculation of material properties	NM coarse-grained presentation, NM descriptors, Hamaker constants	CSV*	NanoCommons, Zenodo
AOP	Statistical Modelling	Computer simulation (CG) of the interaction of the key molecule with the NM	Programs & Computational models	Mathematica NB, ISA-TAB, JSON	SNT, Zenodo
NMs	Molecular simulations	Multiscale modelling of NM corona formation & validation of the modelling against in vivo studies	Bionano interactions descriptors - adsorption free energies; binding coordination	data sheets CSV*	SNT, Zenodo

NM-protein interaction

Each material is linked to a set of 30 PMFs: amino acids, lipid fragments, sugars. Protein adsorption maps



NM Hydration and NM-lipid interaction

Atomistic MD simulations

surface	Au 100	SiO ₂ Q ²	SiO ₂ Q ³	SiO ₂ Q ^{3a}	SiO ₂ Q ⁴	TiO ₂ anatase 101	TiO ₂ anatase 100	TiO ₂ anatase 001	TiO ₂ rutile 110	TiO ₂ rutile 100	TiO ₂ rutile 001	
H ₂ O	water hydrogen bond density (nm ⁻²)	n/a	5.7 ± 0.3	8.3 ± 0.3	6.5 ± 0.3	1.2 ± 0.23	6.3 ± 0.2	8.8 ± 0.3	5.1 ± 0.3	5.9 ± 0.1	10.9 ± 0.5	9.1 ± 0.5
	immersion enthalpy (mN.m ⁻¹)	n/a	-76.7 ± 0.8	-151 ± 1.5	-117 ± 2	22.2 ± 0.8	-774 ± 1	-1176 ± 1	-542 ± 1	-1024 ± 1	-1009 ± 1	-1149 ± 1
	adsorption energy (kJ.mol ⁻¹)	-262.9	-2.5	-15.8	-3.8	-53.6	-2.3	-3.3	n/a	-2.2	-2.4	n/a
	adsorption enthalpy (kJ.mol ⁻¹)	n/a	9.1 ± 8.3	-21 ± 6.9	-69 ± 11	-64.8 ± 9.1	n/a	n/a	n/a	n/a	n/a	n/a
lipid	area per lipid (nm ⁻²)	0.84	0.65	0.67	0.71	0.64	0.65	0.65	0.62	n/a	n/a	n/a
	separation (nm)	0.36	0.76	0.54	0.52	1.01	0.87	0.84	1.15	n/a	n/a	n/a
	thickness (nm)	2.83	3.15	3.08	2.98	3.18	3.04	3.07	3.20	n/a	n/a	n/a
	adhesion strength (mN.m ⁻¹)	-42.2 ± 2	-0.8 ± 0.3	-1.2 ± 0.3	-2.7 ± 0.5	-0.7 ± 0.2	-3.3 ± 0.3	-6.6 ± 0.7	0	>0	>0	>0
	adhesion enthalpy (mN.m ⁻¹)	-118 ± 7.4	2.1 ± 1.2	-2.8 ± 1.7	-10 ± 1.1	-3.4 ± 2.1	-5.5 ± 1.3	n/a	n/a	n/a	n/a	n/a

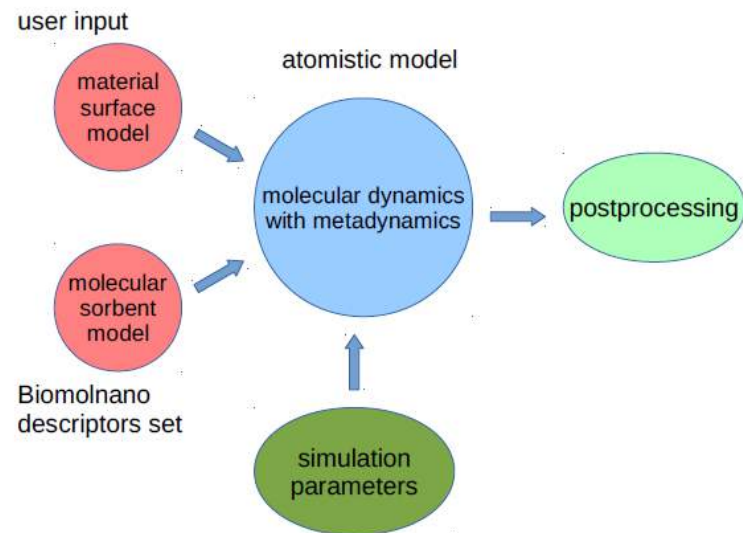
MODA Workflow: bionano interaction PMF

in certain class of solvers for the models.



3 SPECIFIC COMPUTATIONAL MODELLING METADATA		
3.1	NUMERICAL SOLVER	Verlet algorithm with V-rescale thermostat Well-tempered metadynamics algorithm with collective variable determined as the surface separation distance between the sorbent molecule and surface
3.2	SOFTWARE TOOL	Gromacs v. 5.1.2 with Plumed 2.3
3.3	TIME STEP	2 fs, total simulation time ~300 ns
3.4	COMPUTATIONAL REPRESENTATION <i>Refers to how your computational solver represents the material, properties, equation variables,</i>	PHYSICS EQUATION, MATERIAL RELATIONS, MATERIAL Newton's equations are discretized in a Verlet scheme. Forces are computed by summation over all atom pairs within the cut-off radius. Long-range electrostatic interactions are treated by the Particle Mesh Ewald (PME) Well-tempered metadynamics with adaptive Gaussian width
3.5	COMPUTATIONAL BOUNDARY CONDITIONS	Material slab periodic in X and Y direction Water filling the rest of the simulation box
3.6	ADDITIONAL SOLVER PARAMETERS	thermostat relaxation time: 1 ps real space cutoff: 1 nm PME grid spacing: 0.1 nm preequilibration time: 1 ns initial height of Gaussians: 2.5 kJ/nm bias factor: 15 Additional potential wall for the sorbent molecule $U_{wall}(s) = k(s-a)^4$ with the force constant $k = 40 \text{ kJ}/(\text{nm})^4$
3.7	PUBLICATIONS	E. G. Brandt and A. P. Lyubartsev "Molecular Dynamics Simulation of the Interaction of an Amino Acid Side Chain Analogues with a Surface" J. Phys. Chem. C, 119(32), 2015

Workflow picture



Challenges

- Absence of ontologies for novel in vitro, in silico methods and outcomes
- Non-conventional methods and data types

Semantic mapping simulation ↔ experiment

- Level 1: bulk material (chemistry + crystalline form) → NM
- Level 2: nanomaterial model (L1 + specific size + basic extrinsic properties (charge) → NM
- Level 3: nanomaterial model (L2 + size distribution + extrinsic properties (coating)) → NM

Further actions

July-September 2020

- Data organisation and storage – SNT server, GEO, Zenodo
- In vivo/ in vitro data conversion and annotation: TA NanoCommons
- In silico data: data organisation and storage, semantic mapping – NanoCommons

2020-2021

- Models: integration into NanoCommons KB, NanoSolveIT KB, Biovia QSAR WorkBench, Biovia Materials Studio
- All data release to public