# Develop Externally Validated QNAR Models That Can be Reliabliy Used to Prioritize Nanoparticles for Biological and Safety Studies *(425.045)*

## *Development of Novel Descriptors*

**PIs:**

- **Alexander Tropsha, PhD (University of North Carolina at Chapel Hill)**
- **Denis Fourches, PhD (University of North Carolina at Chapel Hill)**

# Research Objectives

- **Develop predictive QNAR models that correlate the compositional/physical/chemical/geometrical and biological descriptors of MNPs with known toxicological endpoints.**

- **Employ QNAR models for virtual screening of libraries of compound considered attachable to CNT surfaces to prioritize compounds for the experimental validation of predicted cellular toxicity and protein binding.**

- **In joint studies with experimental collaborators, evaluate the accuracy of QNAR models by testing the selected MNPs in biological experiments.**

# ESH Metrics and Impact

1. **Obtain reliable descriptors of the physical and chemical properties of manufactured nanoparticles.**

2. **Develop predictive computational models that correlate physical-chemical descriptors of MNPs with their toxic effects.**

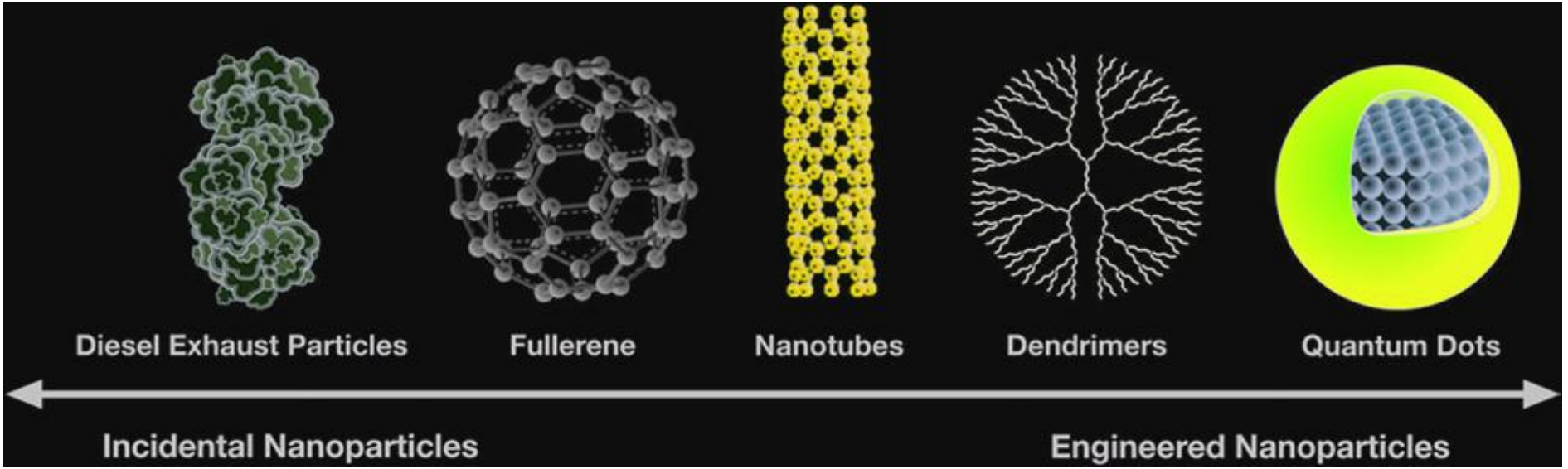
**Impact: Utilize the knowledge gained through above metrics for improved MNP experimental design and prioritized toxicity testing toward the manufacturing of safe nanomaterials.**

# Research Hypothesis

- **The biological/toxicological properties of MNPs (exemplified by carbon nanotubes, CNTs) depend on the compositional/physical/chemical/ geometrical properties of the CNTs.**

- **Using physical/chemical characterization and toxicological screens for an ensemble of MNPs, it will be possible to develop and experimentally validate predictive Quantitative Nanostructure – Activity (QNAR) models.**

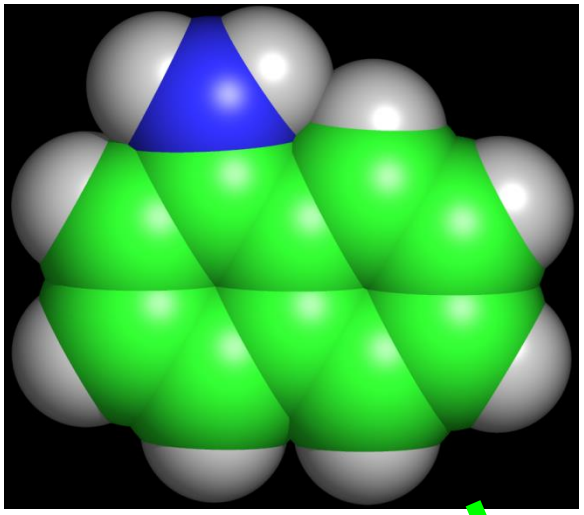
# Challenges in Modeling of Nanomaterials



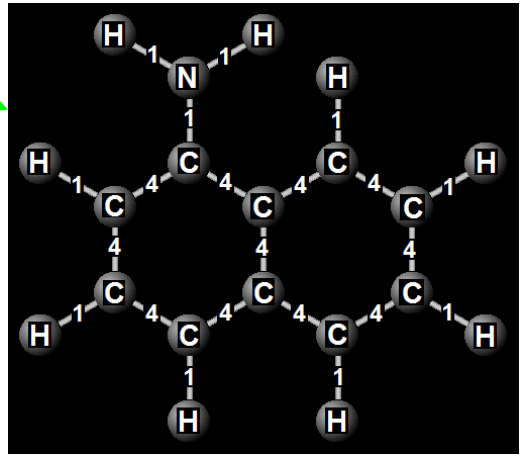

S. Stern and S. McNeil, Toxicological Sciences, 101(1), 4-21, 2008.

- **NP structures are very diverse → a real challenge to develop quantitative parameters (descriptors) of MNPs.**
- **Systematic physico-chemical, geometrical, structural and biological studies of NPs are nearly absent.**
- **Computational modeling of nanoparticles is only beginning to emerge; best if done in collaboration with experimental scientists.**

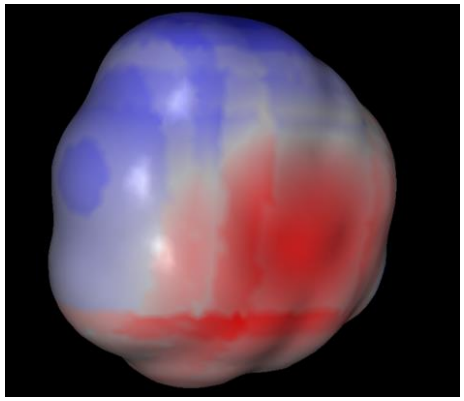
# Structure Representation in Cheminformatics



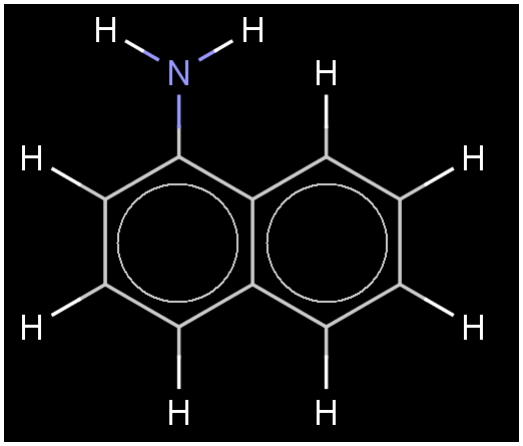
naphtalen-1-amine



Viewed by computers



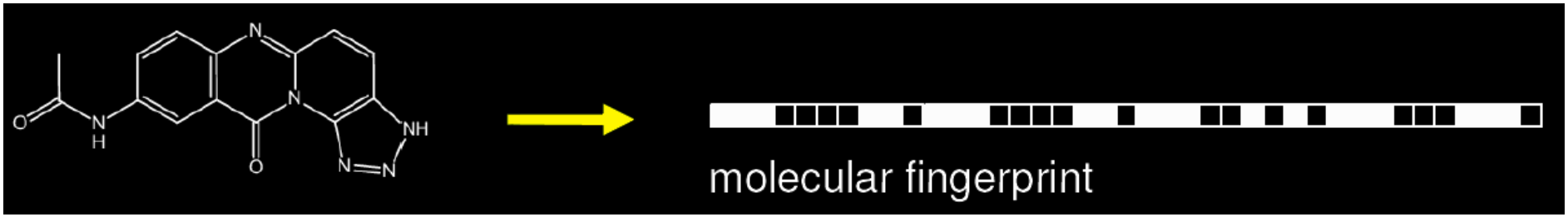
Viewed by another molecule

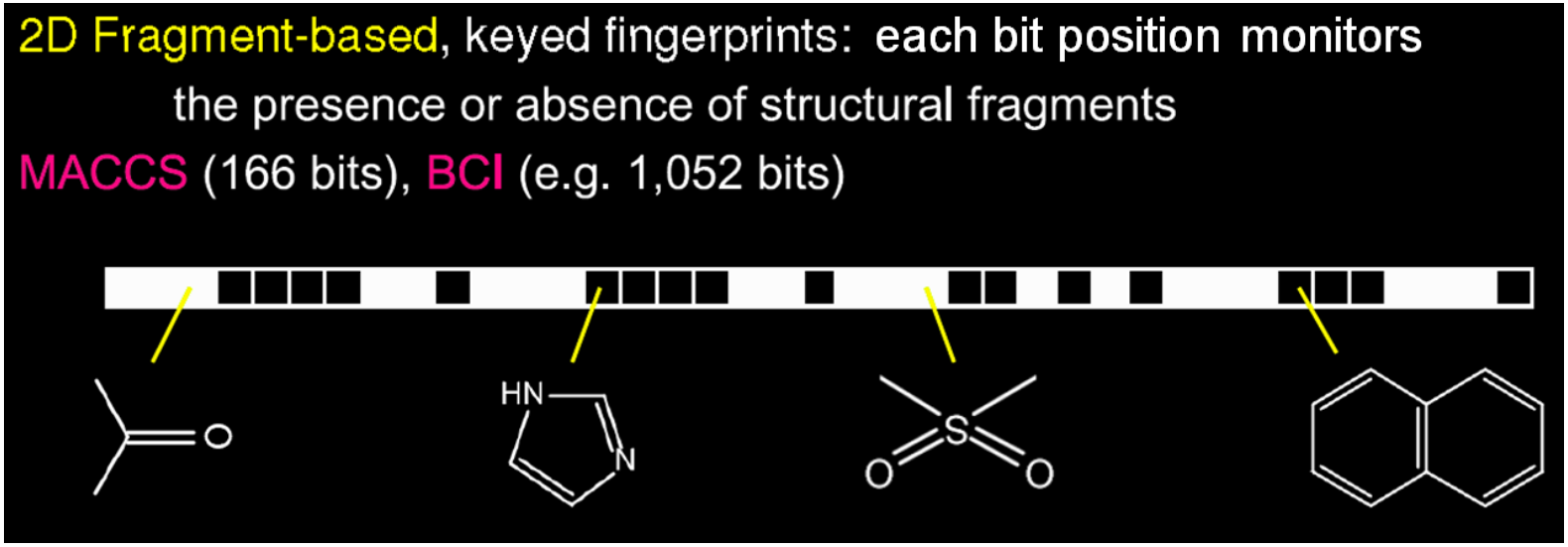


Viewed by chemists

# Molecular Fingerprints

**bit string encodings of**
**structural features and/or calculated molecular properties.**



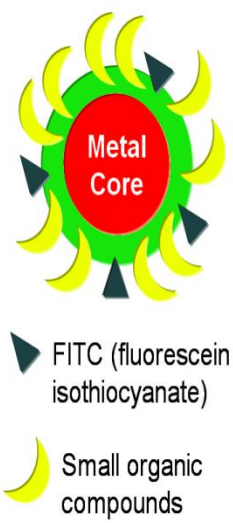
NANOPARTICLES

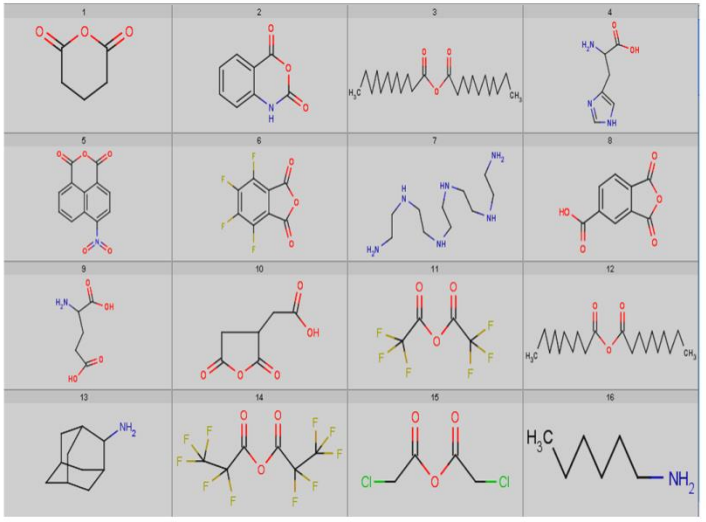
DESCRIPTORS

Computational Descriptors

Experimental properties

**ALL PARTICLES HAVE THE SAME CORE BUT DIFFERENT SURFACE MODIFIERS**

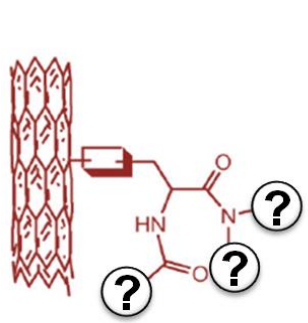


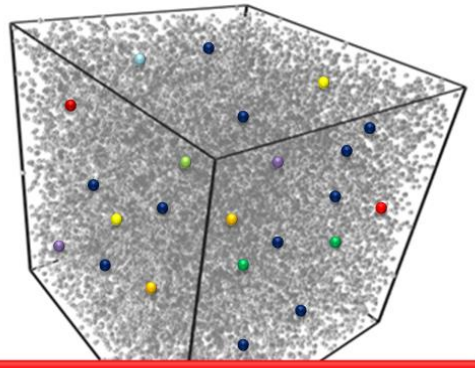



**Classical molecular descriptors (e.g., Dragon, MOE, SiRMS) can be computed for a single molecule that represents the surface of a particular nanoparticle.**

# Computer-aided design of novel carbon nanotubes with desired biological properties

(in collaboration with Dr. Bing Yan, St. Jude Children's Research Hospital)



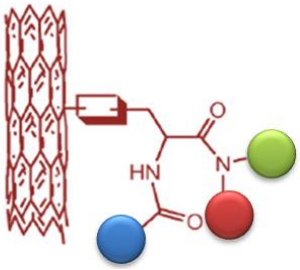


~240,000 small molecules considered attachable to CNTs

QNAR models
Similarity filters
Empirical rules

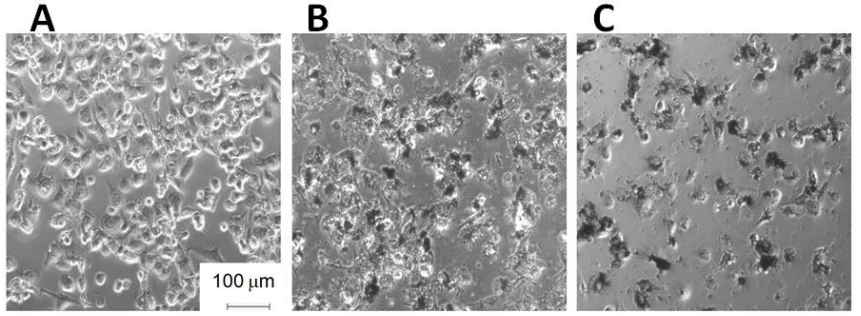
VIRTUAL SCREENING

~$10^2$ molecules



Synthesis and experimental tests

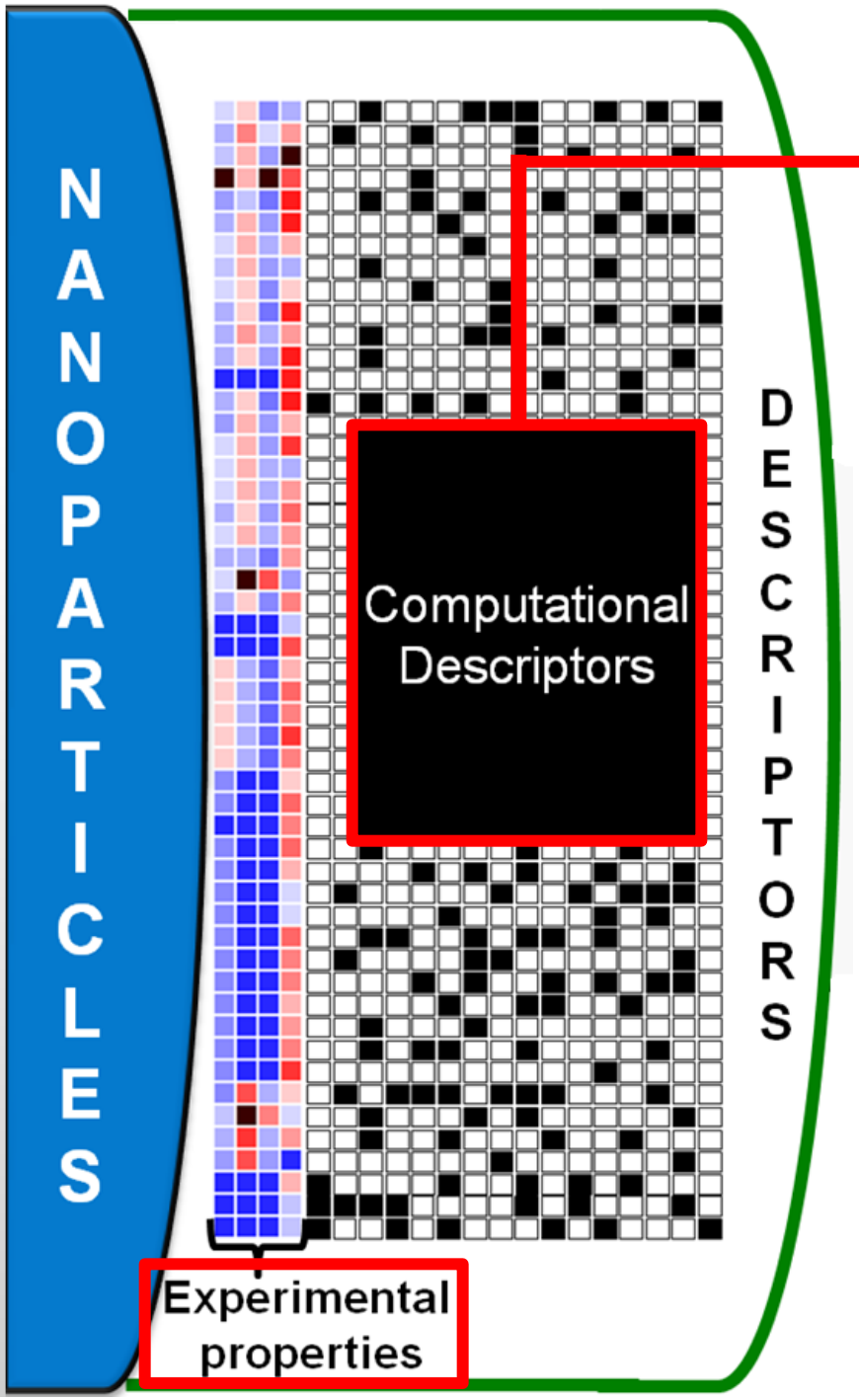
**Excellent agreement between experimental and predicted activities**




D

| CNT ID | II-1 | II-2 | II-3 | II-4 | II-5 | II-6 | II-7 | II-8 | II-9 | II-10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Average cell viability (%) | 58 | 61 | 61 | 56 | 58 | 65 | 68 | 72 | 68 | 59 |
| Standard Deviation (%) | 5 | 3 | 3 | 3 | 2 | 10 | 6 | 7 | 3 | 6 |
| Experiment | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Predicted | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| CNT ID | II-11 | II-12 | II-13 | II-14 | II-15 | II-16 | II-17 | II-18 | II-19 | II-20 |
| Average cell viability (%) | 39 | 49 | 46 | 49 | 52 | 41 | 51 | 49 | 55 | 50 |
| Standard Deviation (%) | 9 | 8 | 7 | 5 | 8 | 11 | 5 | 9 | 11 | 10 |
| Experiment | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| Predicted | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

| CNT ID | II-21 | II-22 | II-23 | II-24 | II-25 | II-26 | II-27 | II-28 | II-29 | II-30 |
|---|---|---|---|---|---|---|---|---|---|---|
| Average protein binding (F0/F1) | 1.77 | 1.78 | 1.87 | 1.76 | 1.82 | 2.30 | 1.74 | 2.00 | 1.67 | 2.33 |
| Standard Deviation | 0.05 | 0.06 | 0.02 | 0.03 | 0.02 | 0.02 | 0.02 | 0.01 | 0.06 | 0.02 |
| Experiment | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 |
| Predicted | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| CNT ID | II-31 | II-32 | II-33 | II-34 | II-35 | II-36 | II-37 | II-38 | II-39 | II-9 |
| Average protein binding (F0/F1) | 3.40 | 2.28 | 2.04 | 2.22 | 2.17 | 2.95 | 2.08 | 2.25 | 2.65 | 2.24 |
| Standard Deviation | 0.03 | 0.05 | 0.08 | 0.01 | 0.04 | 0.00 | 0.02 | 0.06 | 0.05 | 0.11 |
| Experiment | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Predicted | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |




**ALL PARTICLES HAVE DIFFERENT CORES/ARCHITECTURES**

| MNP | CLIO | PNP | MION | QD | Feridex IV | Ferrum Hausmann |
|---|---|---|---|---|---|---|
| #. particle | 23 | 19 | 4 | 3 | 1 | 1 |

*Shaw et al., PNAS, 2008, 105, 7387-7392*

If no available three-dimensional structures, only constitutional descriptors (e.g., number of metal atoms, presence/ absence of coated dextran) are computationally accessible.

**Need of developing new descriptors**

**Development of Quantum-Mechanics based fingerprints in collaboration with Dr. Stefano Curtarolo *(Duke University)***

# New theoretical MNP descriptors: Computed characteristics of materials*

| Index | ▲ Name | Atom Number | Fit Band Gap (eV) | me(m0) | mmine(m0) | mh(m0) | mminh(m0) | Mass ratio | Valence Band Width (eV) | Core Valence Gap (eV) | Density (g/cm3) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Al1La1O3 | 5 | 5.13 | 0.45 | 0.36 | 1.81 | 0.56 | 4.06 | 7.46 | 6.06 | 6.42 |
| 2 | As1Ca3Cl3 | 7 | 3.39 | 0.64 | 0.34 | 0.52 | 0.24 | 1.24 | 1.36 | 1.62 | 2.63 |
| 3 | As1Cl1Hg3O4 | 9 | 3.67 | 1.21 | 1.21 | 2.6 | 2.42 | 2.15 | 6.21 | 0.18 | 8.76 |
| 4 | Ba1Ce1O3 | 5 | 4.9 | 2.54 | 1.81 | 3.69 | 0.3 | 1.45 | 3.15 | 8.17 | 6.44 |
| 5 | Ba1F3Li1 | 5 | 9.68 | 0.63 | 0.51 | 12.87 | 0.8 | 20.52 | 2.31 | 6.27 | 5.25 |
| 6 | Ba1O3Pr1 | 5 | 3.63 | 2.9 | 1.9 | 4.75 | 0.3 | 1.63 | 3.07 | 8.07 | 6.52 |
| 7 | Ba1O3Zr1 | 5 | 5.01 | 0.69 | 0.37 | 2.92 | 0.9 | 4.2 | 3.64 | 6.75 | 6.28 |
| 8 | Ba1Se1 | 2 | 2.55 | 0.55 | 0.23 | 0.67 | 0.19 | 1.23 | 3.06 | 7.49 | 6.57 |
| 9 | Be1H3Na1 | 5 | 2.15 | 0.51 | 0.49 | 0.41 | 0.23 | 1.25 | 9.63 | 14.17 | 1.67 |
| 10 | Bi1In1O3 | 5 | 1.22 | 0.26 | 0.22 | 11.71 | 0.35 | 44.78 | 7.13 | 2.87 | 8.37 |
| 11 | Br1K3O1 | 5 | 2.11 | 0.38 | 0.37 | 2.35 | 0.55 | 6.27 | 0.92 | 2.62 | 2.5 |
| 12 | Br1O1Rb3 | 5 | 1.46 | 0.34 | 0.33 | 2.13 | 0.41 | 6.35 | 1.07 | 2.66 | 3.58 |
| 13 | Br3Cd1Cs1 | 5 | 2.35 | 0.19 | 0.18 | 6.39 | 0.28 | 34.39 | 5.55 | 0.78 | 5.32 |
| 14 | C1K4O4 | 9 | 3.96 | 0.5 | 0.5 | 8.74 | 6.45 | 17.55 | 0.13 | 0.69 | 2.65 |
| 15 | C1Li4O4 | 9 | 6.53 | 0.8 | 0.79 | 3.45 | 1.84 | 4.28 | 4.28 | 1.66 | 2.63 |
| 16 | C1Na4O4 | 9 | 3.67 | 0.48 | 0.48 | 5.38 | 3.89 | 11.21 | 0.68 | 0.33 | 2.97 |

*data from aflowlib.org (collaboration with Prof. Stefano Curtarolo, Duke University)

# New theoretical MNP descriptors: Materials fingerprints

| Index | Name | Atom Number | Fit Band Gap (eV) | me(m0) | mmine(m0) | mh(m0) | mminh(m0) | Mass ratio | Valence Band Width (eV) | Core Valence Gap (eV) | Density (g/cm3) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Al1La1O3 | 5 | 5.13 | 0.45 | 0.36 | 1.81 | 0.56 | 4.06 | 7.46 | 6.06 | 6.42 |
| 2 | As1Ca3Cl3 | 7 | 3.39 | 0.64 | 0.34 | 0.52 | 0.24 | 1.24 | 1.36 | 1.62 | 2.63 |
| 3 | As1Cl3O4 | 9 | 3.67 | 1.21 | 1.21 | 2.6 | 2.42 | 2.15 | 6.21 | 0.18 | 8.76 |
| 4 | Ba1Cl1O3 | 5 | 4.9 | 2.54 | 1.81 | 3.69 | 0.3 | 1.45 | 3.15 | 8.17 | 6.44 |
| 5 | Ba1I1Li1 | | | | | | | | | | |
| 6 | Ba1Cl1Pr1 | | | | | | | | | | |
| 7 | Ba1O3Zr1 | | | | | | | | | | |
| 8 | Ba1Se1 | | | | | | | | | | |
| 9 | Be1H3Na1 | 5 | 2.15 | 0.51 | 0.49 | 0.41 | 0.23 | 1.25 | 9.63 | 14.17 | 1.67 |
| 10 | Bi1In1O3 | 5 | 1.22 | 0.26 | 0.22 | 11.71 | 0.35 | 44.78 | 7.13 | 2.87 | 8.37 |
| 11 | Br1K3O1 | 5 | 2.11 | 0.38 | 0.37 | 2.35 | 0.55 | 6.27 | 0.92 | 2.62 | 2.5 |
| 12 | Br1O1Rb3 | 5 | 1.46 | 0.34 | 0.33 | 2.13 | 0.41 | 6.35 | 1.07 | 2.66 | 3.58 |
| 13 | Br3Cd1Cs1 | 5 | 2.35 | 0.19 | 0.18 | 6.39 | 0.28 | 34.39 | 5.55 | 0.78 | 5.32 |
| 14 | C1K4O4 | 9 | 3.96 | 0.5 | 0.5 | 8.74 | 6.45 | 17.55 | 0.13 | 0.69 | 2.65 |
| 15 | C1Li4O4 | 9 | 6.53 | 0.8 | 0.79 | 3.45 | 1.84 | 4.28 | 4.28 | 1.66 | 2.63 |
| 16 | C1Na4O4 | 9 | 3.67 | 0.48 | 0.48 | 5.38 | 3.89 | 11.21 | 0.68 | 0.33 | 2.97 |
| 17 | C2Ca1 | 3 | 3.28 | 1.33 | 0.33 | 1.56 | 0.59 | 1.17 | 1.81 | 0.49 | 2.25 |
| 18 | Ca1O3Ti1 | 5 | 4.1 | 1 | 0.45 | 2.66 | 0.82 | 2.65 | | | |
| 19 | Ca1O3Zr1 | 5 | 5.25 | 0.7 | 0.39 | 3.09 | 0.93 | 4.39 | | | |
| 20 | Ca3Cl3P1 | 7 | 3.4 | 0.63 | 0.34 | 0.56 | 0.27 | 1.13 | | | |
| 21 | Cd1F3Rb1 | 5 | 5.58 | 0.44 | 0.44 | 40.03 | 0.52 | | | | |
| 22 | Cl1Rb1 | 2 | 7.33 | 0.39 | 0.39 | 3.31 | 0.7 | | | | |
| 23 | Cl3Cs1Sn1 | 5 | 1.75 | 0.47 | 0.08 | 0.11 | 0.1 | | | | |
| 24 | Cr1La1O3 | 5 | 3.33 | 0.72 | 0.36 | 1.69 | 1.45 | | | | |
| 25 | Cs1F3Hg1 | 5 | 1.89 | 0.34 | 0.33 | 325.67 | 0.43 | 968.99 | | | |
| 26 | Cs1F3Mg1 | 5 | 9.98 | 0.45 | 0.44 | 7.73 | 0.54 | 17.32 | | | |
| 27 | Cs3O3Sb1 | 7 | 5.18 | 1.96 | 0.32 | 2.23 | 1.63 | 1.14 | | | |
| 28 | Cs3Sb1Se3 | 7 | 4.05 | 5.49 | 1.76 | 3.68 | 0.9 | 1.49 | | | |
| 29 | Cu2O1 | 3 | 1.79 | 0.82 | 0.82 | 1.87 | 0.22 | 2.28 | 6.84 | 11.98 | 6.14 |
| 30 | Cu3S4Ta1 | 8 | 3.84 | 1.31 | 0.97 | 1.24 | 0.62 | 1.06 | 5.3 | 6.82 | 4.93 |
| 31 | F3Mg1Na1 | 5 | 8.9 | 0.56 | 0.56 | 11.78 | 1.83 | 21.02 | 3.94 | 15.54 | 2.97 |
| 32 | F3Rb1V1 | 5 | 3.86 | 0.49 | 0.49 | 12.81 | 1.98 | 26.01 | 7.49 | 3.25 | 4.39 |
| 33 | F3Rb1Yb1 | 5 | 2.39 | 0.49 | 0.49 | 9.59 | 3.96 | 19.35 | 0.19 | 4.86 | 5.63 |
| 34 | F3Sc1 | 4 | 9.1 | 4.35 | 1.17 | 23.29 | 2.35 | 5.36 | 3.03 | 16.02 | 2.62 |
| 35 | Fe1La1O3 | 5 | 1.63 | 1.11 | 0.53 | 11.35 | 0.36 | 10.23 | 9.02 | 6.58 | 6.85 |
| 36 | Ge1I3Rb1 | 5 | 1.66 | 0.38 | 0.06 | 0.12 | 0.1 | 3.12 | 4.43 | 2.4 | 4.15 |
| 37 | I1K3O1 | 5 | | | | | | | | | |
| 38 | Ir1S1Sb1 | 3 | | | | | | | | | |
| 39 | N2O1 | 3 | | | | | | | | | |
| 40 | O2Si1 | 3 | | | | | | | | | |
| 41 | O3Pb1Ti1 | 5 | | | | | | | | | |
| 42 | O3Sn1Sr1 | 5 | | | | | | | | | |
| 43 | O3Sr1Tc1 | 5 | | | | | | | | | |
| 44 | O3Sr1Ti1 | 5 | | | | | | | | | |
| 45 | O6Se2Sn1 | 9 | | | | | | | | | |
| 46 | Os1Se2 | 3 | | | | | | | | | |

Etc.

46 materials (*e.g.*, $CaO_3Zr$, BaSe, etc.)

Data Curation Normalization and Selection of Descriptors

## Materials Fingerprints

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0.429 | 0.456 | 0.044 | 0.129 | 0.005 | 0.072 | 0.003 | 0.732 | 0.376 | 0.515 |
| 0.714 | 0.260 | 0.076 | 0.121 | 0.001 | 0.022 | 0.000 | 0.123 | 0.098 | 0.113 |
| 1.000 | 0.292 | 0.173 | 0.496 | 0.008 | 0.365 | 0.001 | 0.607 | 0.008 | 0.763 |
| 0.429 | 0.430 | 0.398 | 0.754 | 0.011 | 0.032 | 0.000 | 0.301 | 0.508 | 0.517 |
| 0.429 | 0.966 | 0.075 | 0.194 | 0.039 | 0.110 | 0.020 | 0.218 | 0.389 | 0.391 |
| 0.429 | 0.287 | 0.459 | 0.793 | 0.014 | 0.032 | 0.001 | 0.293 | 0.502 | 0.525 |
| 0.429 | 0.442 | 0.085 | 0.134 | 0.009 | 0.126 | 0.003 | 0.350 | 0.419 | 0.500 |
| 0.000 | 0.166 | 0.061 | 0.073 | 0.002 | 0.014 | 0.000 | 0.292 | 0.466 | 0.531 |
| 0.429 | 0.121 | 0.054 | 0.185 | 0.001 | 0.021 | 0.000 | 0.948 | 0.884 | 0.012 |
| 0.429 | 0.017 | 0.012 | 0.069 | 0.036 | 0.039 | 0.045 | 0.699 | 0.176 | 0.721 |
| 0.429 | 0.117 | 0.032 | 0.134 | 0.007 | 0.071 | 0.005 | 0.079 | 0.160 | 0.100 |
| 0.429 | 0.044 | 0.025 | 0.116 | 0.006 | 0.049 | 0.006 | 0.094 | 0.163 | 0.214 |
| 0.429 | 0.144 | 0.000 | 0.052 | 0.019 | 0.028 | 0.034 | 0.541 | 0.045 | 0.398 |
| 1.000 | 0.324 | 0.053 | 0.190 | 0.027 | 1.000 | 0.017 | 0.000 | 0.040 | 0.116 |
| 1.000 | 0.613 | 0.103 | 0.315 | 0.010 | 0.274 | 0.003 | 0.414 | 0.100 | 0.113 |
| 1.000 | 0.292 | 0.049 | 0.181 | 0.016 | 0.597 | 0.011 | 0.055 | 0.017 | 0.149 |
| 0.143 | 0.248 | 0.193 | 0.116 | 0.005 | 0.077 | 0.000 | 0.168 | 0.027 | 0.073 |
| 0.429 | 0.340 | 0.137 | 0.168 | 0.008 | 0.113 | 0.002 | 0.468 | 0.702 | 0.272 |
| 0.429 | 0.469 | 0.086 | 0.142 | 0.009 | 0.131 | 0.003 | 0.395 | 0.689 | 0.320 |
| 0.714 | .262 | 0.075 | 0.121 | 0.001 | 0.027 | 0.000 | 0.117 | 0.103 | 0.082 |
| | | .06 | 0.042 | 0.164 | 0.123 | 0.066 | 0.092 | 0.412 | 0.063 | 0.361 |
| | | | 0.034 | 0.142 | 0.010 | 0.095 | 0.008 | 0.128 | 0.439 | 0.238 |
| | | | 0.047 | 0.009 | 0.000 | 0.000 | 0.003 | 0.589 | 0.000 | 0.212 |
| | | .94 | 0.090 | 0.129 | 0.005 | 0.213 | 0.001 | 0.604 | 0.540 | 0.543 |
| 0.429 | | .092 | 0.025 | 0.116 | 1.000 | 0.052 | 1.000 | 0.606 | 0.788 | 0.554 |
| 0.429 | | 1.000 | 0.044 | 0.164 | 0.023 | 0.069 | 0.017 | 0.476 | 0.898 | 0.333 |
| 0.714 | | 0.461 | 0.300 | 0.112 | 0.007 | 0.241 | 0.000 | 0.021 | 0.013 | 0.359 |
| 0.714 | | 0.335 | 0.897 | 0.733 | 0.011 | 0.126 | 0.000 | 0.001 | 0.002 | 0.283 |
| 0.143 | | 0.081 | 0.107 | 0.328 | 0.005 | 0.019 | 0.001 | 0.670 | 0.747 | 0.485 |
| 0.857 | | 0.311 | 0.190 | 0.392 | 0.004 | 0.082 | 0.000 | 0.516 | 0.424 | 0.357 |
| 0.429 | | 0.879 | 0.063 | 0.216 | 0.036 | 0.272 | 0.021 | 0.380 | 0.970 | 0.149 |
| 0.429 | | 0.313 | 0.051 | 0.185 | 0.039 | 0.296 | 0.026 | 0.735 | 0.200 | 0.300 |
| 0.429 | | 0.148 | 0.051 | 0.185 | 0.029 | 0.608 | 0.019 | 0.006 | 0.301 | 0.431 |
| 0.286 | | 0.901 | 0.704 | 0.478 | 0.071 | 0.354 | 0.004 | 0.289 | 1.000 | 0.112 |
| 0.429 | | 0.063 | 0.156 | 0.203 | 0.035 | 0.041 | 0.010 | 0.887 | 0.409 | 0.560 |
| 0.429 | | 0.066 | 0.032 | 0.000 | 0.000 | 0.000 | 0.002 | 0.429 | 0.147 | 0.274 |
| 0.429 | | 0.137 | 0.025 | 0.121 | 0.007 | 0.074 | 0.006 | 0.072 | 0.150 | 0.145 |
| 0.143 | | 0.199 | 0.164 | 0.392 | 0.001 | 0.057 | 0.002 | 0.594 | 0.149 | 0.951 |
| 0.143 | | 0.487 | 0.880 | 1.000 | 0.013 | 0.370 | 0.000 | 0.067 | 0.246 | 0.000 |
| 0.143 | | 0.589 | 0.029 | 0.125 | 0.009 | 0.088 | 0.008 | 1.000 | 0.424 | 0.321 |
| 0.429 | | 0.312 | 1.000 | 0.207 | 0.002 | 0.025 | 0.007 | 0.487 | 0.035 | 0.686 |
| 0.429 | | 0.125 | 0.007 | 0.060 | 0.009 | 0.098 | 0.012 | 0.877 | 0.367 | 0.520 |
| 0.429 | | 0.000 | 0.117 | 0.108 | 0.004 | 0.145 | 0.001 | 0.691 | 0.558 | 0.504 |
| 0.429 | | 0.329 | 0.149 | 0.172 | 0.008 | 0.112 | 0.002 | 0.460 | 0.602 | 0.377 |
| 1.000 | | 0.370 | 0.036 | 0.147 | 0.018 | 0.652 | 0.014 | 0.271 | 0.071 | 0.315 |
| 0.143 | | 0.019 | 0.014 | 0.086 | 0.003 | 0.153 | 0.003 | 0.622 | 0.313 | 1.000 |

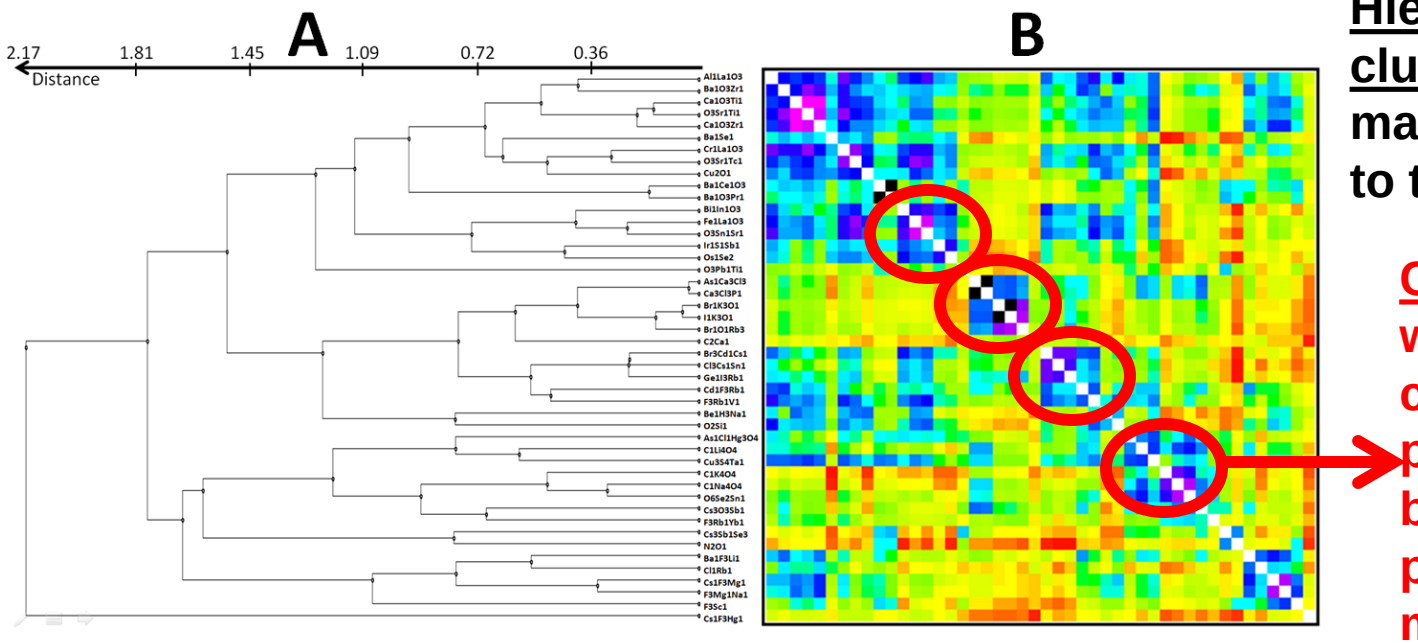
Descriptors derived from quantum mechanics calculations (DFT)
*In collaboration with Dr. Curtarolo (Duke University)*

# Materials similarity based on their fingerprints

Tanimoto similarity coefficient S between materials A and B is calculated as follows:

$$S_{A,B} = [\sum_{j=1}^{j=n} x_{jA}x_{jB}]/[\sum_{j=1}^{j=n}(x_{jA})^2 + \sum_{j=1}^{j=n}(x_{jB})^2 - \sum_{j=1}^{j=n} x_{jA}x_{jB}]$$

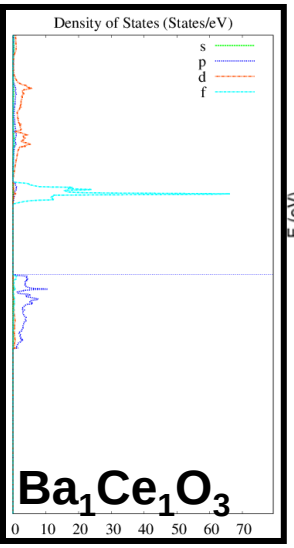
with $x_j$ is the value of the $j^{th}$ descriptor and *n* the total number of descriptors in the fingerprints. Tanimoto similarities are ranging from 0 (no similarity between materials A and B) to 1 (A and B are identical).

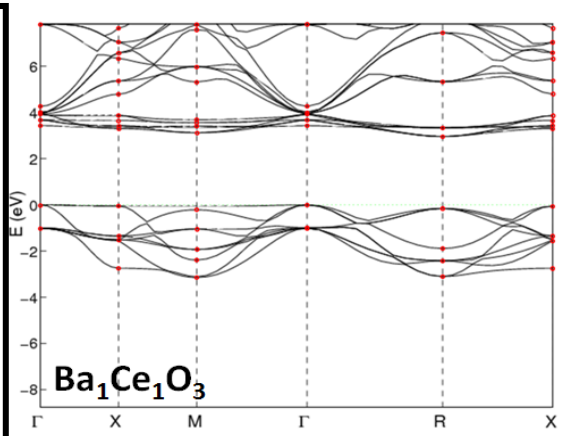


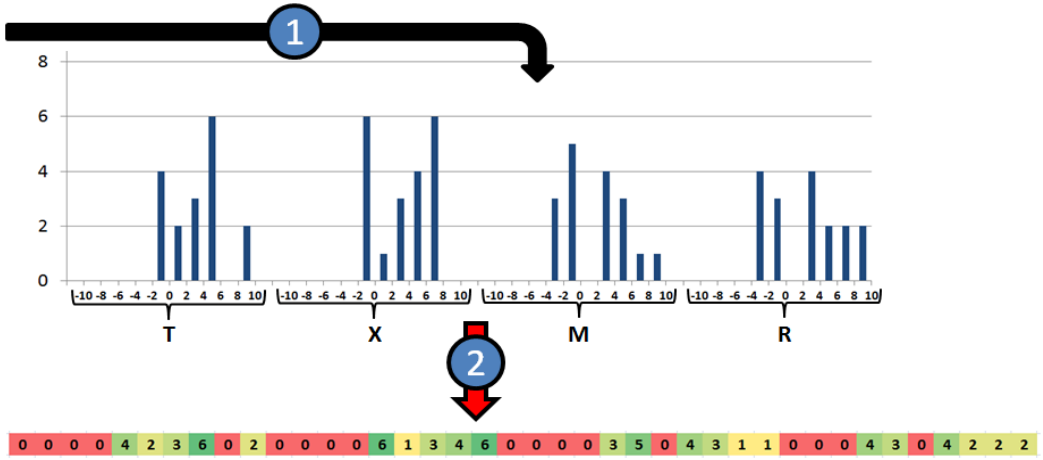

**Hierarchical clustering of 46 materials according to their fingerprints.**

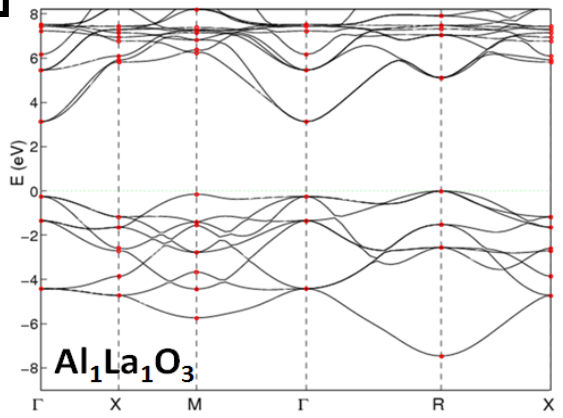
**Cluster Analysis: within a given cluster, analysis of physical/chemical/ biological properties of materials**

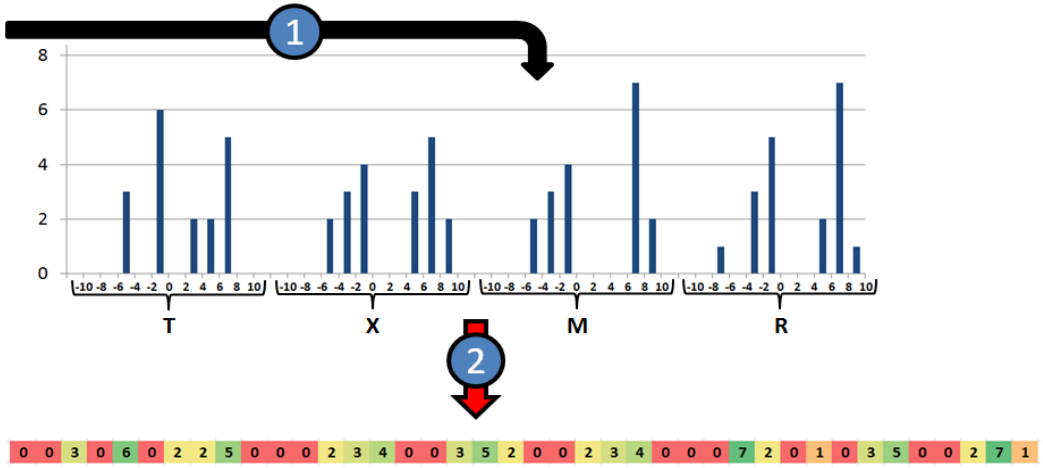
# Material Fingerprints



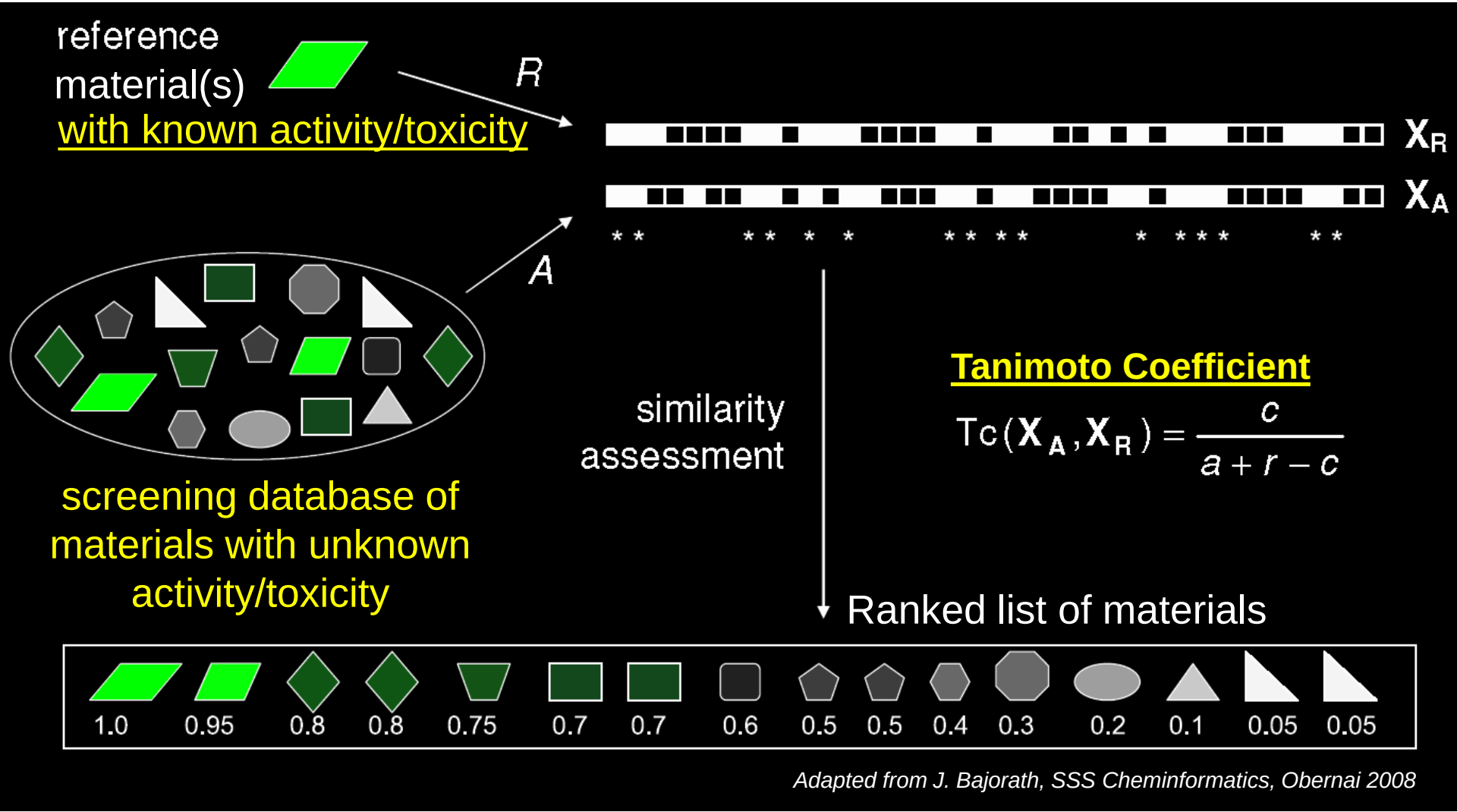




Materials fingerprints generated from electronic band structures computed with DFT. Band structures are transformed into band distribution plots **(1)** and then converted into materials fingerprints **(2)**.

***Our fingerprints can also be computed from material's Density of State.***

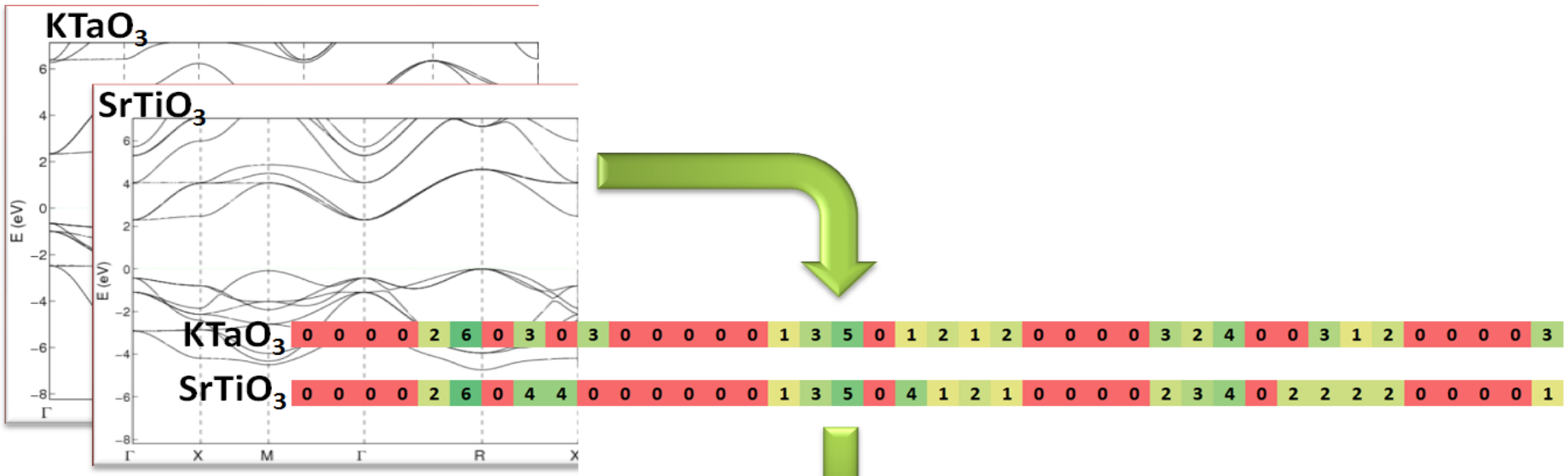
# Similarity Search

**Similarity searching using fingerprint representations of chemicals is one of the most widely used approaches for chemical database mining: it assumes that similar compounds possess similar biological activities.**




*Adapted from J. Bajorath, SSS Cheminformatics, Obernai 2008*

# Reprofiling materials with desired properties




**TANIMOTO_SIMILARITY ($KTaO_3$, $SrTiO_3$) = 0.74**

**BOTH COMPOUNDS ARE SIMILAR BASED ON THEIR BAND STRUCTURES**

*"$KTaO_3$ is a promising candidate for superconductivity induced by electrostatic doping because it is similar to the superconductor $SrTiO_3$: [...] have similar band structures, and both exhibit quantum para-electricity"*

**Ueno et al. Nature Nanotechnology, May 2011, Online Epub.**

# Future Plans

## Next Year Plans

- ***DONE*** **- Calculate Fingerprints for all metal alloys (~17,000 different materials) stored in AFLOWLIB data base;**
- **Rank materials based on their similarity to the most toxic ones based on *in vitro* and *in vivo* experimental tests;**
- **New iteration of design/screening/test of carbon nanotubes with selective protein binding and low acute toxicity.**

## Long-Term Plans

- **Assemble a standard set of MNPs fingerprints to build predictive QNAR models;**
- **Collect all QNAR models and make them freely accessible on our Chembench webportal.**

# Industrial Interactions and Technology Transfer

- **Duke University – Dr. Curtarolo**
  **DFT calculations for metal alloys**

- **Office of Naval Research – Development of Predictive Materials Fingerprints for designing materials with desired properties**

- ***In Progress* – NSF grant proposal for nanomaterials design (our group is looking for industrial partners)**

# Publications, Presentations, and Recognitions/Awards

- **Grant awarded by Office of Naval Research ($125k /year for 3 years)**
- **Fourches et al. Computer-Aided Design of Carbon Nanotubes with the Desired Bioactivity and Safety Profiles. Nature Nanotechnology. *Under Review*.**
- **SOT 2013 – Presentation entitled "Quantitative Nanostructure-Activity Relationships"**
- **MRS 2012 – Presentation entitled "Rational Design of (Nano)materials with the Desired Biological Properties Using Quantitative Structure-property Relationships (QSPR) Modeling"**