# *Virtual High-Throughput Screening: How to Boost the Pharmacophore Approach?*

Frédérique Barbosa
Molecular Modelling
F.Barbosa@cerep.fr

# Virtual Screening

- Ligand-Based
  - 2D or 3D descriptions
- Structure-Based
  - 3D description
  - Scoring functions
- Speed of the screening
- Molecules to be screened

# Why virtual screening?

- Find new ligands for a known target
  - Original pharmacophore
  - Patentable compounds
  - ADME/toxicity issues
  - Exhausted lead optimization
- Virtual approach
  - No need of synthesis
  - Cost reduced
- Pharmacophore approach
  - Switching chemotypes
  - Same properties but new families

# Cerep Approach

- Ligand-Based
  - 3D descriptions
  - Pharmacophore approach
- Virtual Library of 100 million compounds
  - Feasibility
  - Availability
- Speed of the screening
  - Less than a day

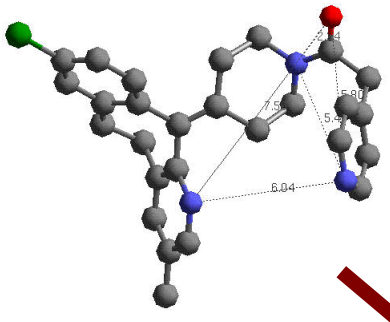
# Finding New Lead Series

- Known synthetic ligands of an enzyme
- Find neighbors in the Potential virtual library
- Analysis of virtual screenings
  - Adding other information as docking, QSAR model, predicted ADME properties, …
  - Choice of best series for synthesis
  - Synthesis of focused libraries
  - Testing

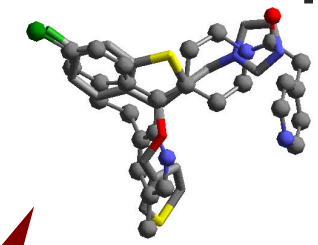
Find new lead series active in the micromolar range or less

# Virtual screening: Theory

Known ligand

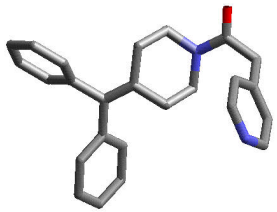


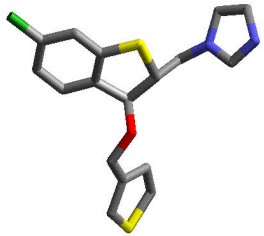
New Active compounds



Descriptor-based Similarity Scoring

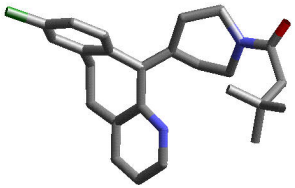
Superposition-based Similarity Scoring

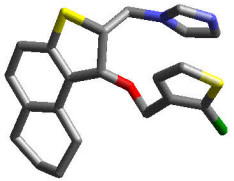




Potential Pharmacophore Fingerprint Library

Best Matching Candidates in terms of FBPA





Cerep

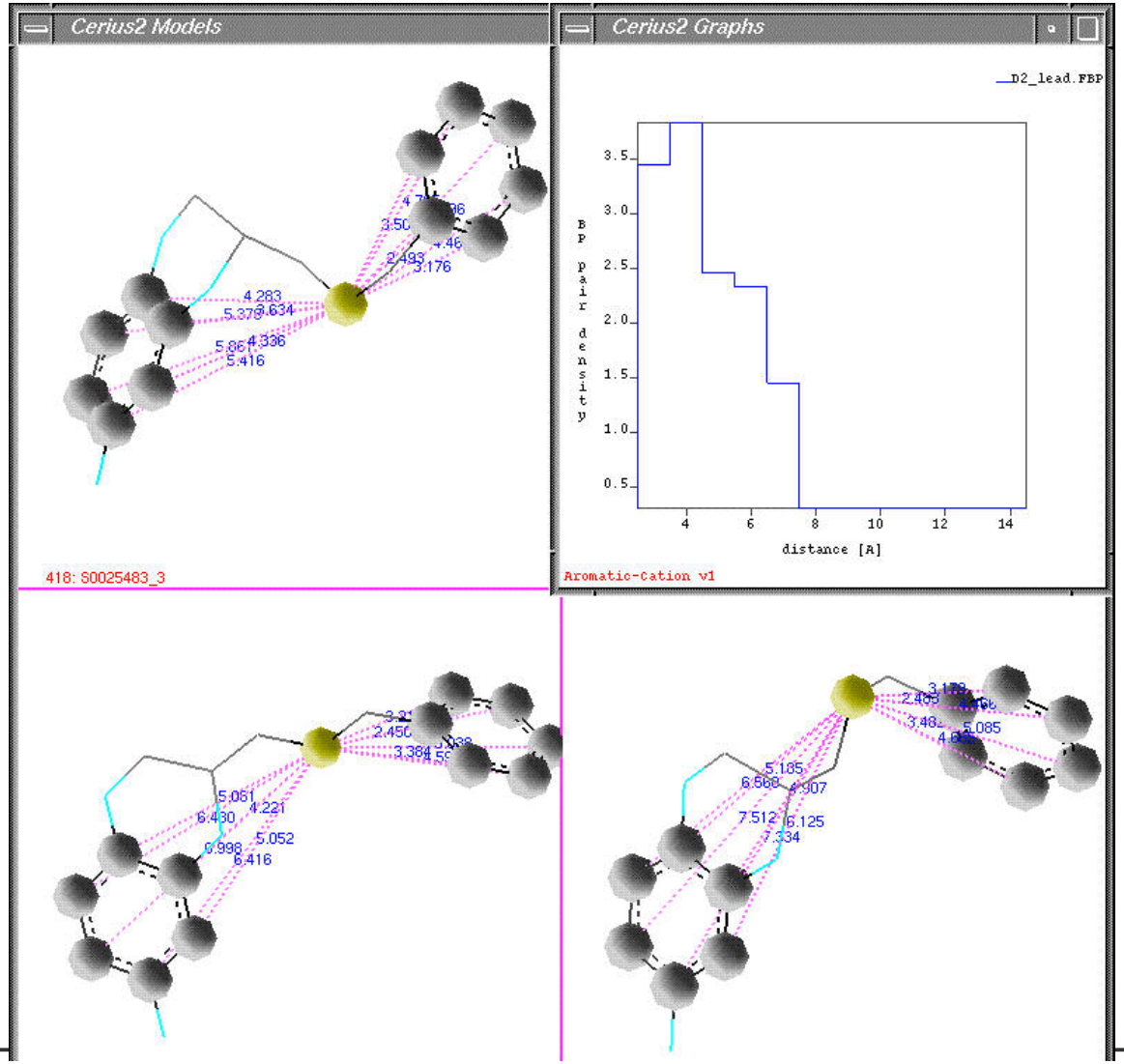
# Fuzzy Bipolar Pharmacophoric Autocorrelograms (FBPA)

- Pharmacophore features:
  - Hydrophobicity
  - Aromaticity
  - HB Acceptor/Donor
  - Positive/Negative Charge
- An atom may possess one or more features
  - Automatic feature recognition routine
- Bipolar Pharmacophores (BP)
  - 21 pairwise combinations of these 6 features: (hydrophobic-hydrophobic), (hydrophobic-aromatic),...(anion-anion)
- Distance Bins
  - 12 considered distance ranges: (3..4), (4..5), …, (14..15) Å

# Fuzzy Bipolar Pharmacophoric Autocorrelograms (FBPA)

- 21 pairsx12 distance bins
  ⇒ 252 descriptors

- FBPA based on multi-conformational models




Cerep

# Similarity metric

- Fuzzy logic in the scoring scheme accounts for the similarity of molecules with atom pairs of same pharmacophore type, when these pairs are assigned to close, albeit different distance bins $\Delta_1 \approx \Delta_2$

$$\left(\Psi_m \quad \Psi_M\right)_{a,b} = \sum_{\Delta_1=1}^{N_{bin}} \sum_{\Delta_2=1}^{N_{bin}} \Psi_m(a,b,\Delta_1)\Psi_M(a,b,\Delta_2)exp\left[-\gamma(\Delta_1-\Delta_2)^2\right]$$

- Feature weighing allows to control the relative importance of each pharmacophore type in the similarity score…

$$\Sigma(m,M) = 1 - \frac{2 \sum_{a,b} w_a w_b \left(\Psi_m \quad \Psi_M\right)_{a,b}}{\sum_{a,b} w_a w_b \left(\Psi_M \quad \Psi_M\right)_{a,b} + \sum_{a,b} w_a w_b \left(\Psi_m \quad \Psi_m\right)_{a,b}}$$

# Behavior of Cerep Metric

- Comparison of diverse metrics
  - Cerep with or without fuzziness
  - Dice
  - Euclidian
- Comparison of diverse descriptors
  - Fuzzy Bipolar Pharmacophoric Autocorrelogram: FBPA
  - 3 point and 4 point pharmacophores
  - 2D descriptors: topological indices, E-state keys
- Criteria for the comparison
  - Consistency
  - Completeness

# Behavior of the metric

- Consistency criterion: $$\chi^{\Sigma}(s) = \frac{\langle \Lambda(M,m) \rangle_P - \langle \Lambda(M,m) \rangle_{P(\Sigma<s)}}{\langle \Lambda(M,m) \rangle_P - \langle \Lambda(M,m) \rangle_{opt}}$$
  - similarity metric void of any NB: $\chi(s)$ 0
  - ideal similarity metric: $\chi(s)$ 1
- Completeness criterion: $$\Omega^{\Sigma}(s) = \frac{\kappa N_{FS} + N_{PFD}}{\kappa N_{FS}^{(rand)} + N_{PFD}^{(rand)}}$$
  - strong NB should translate in minimal numbers of “False Similar” ($N_{FS}$) and “Potentially False Dissimilar” ($N_{PFD}$) pairs

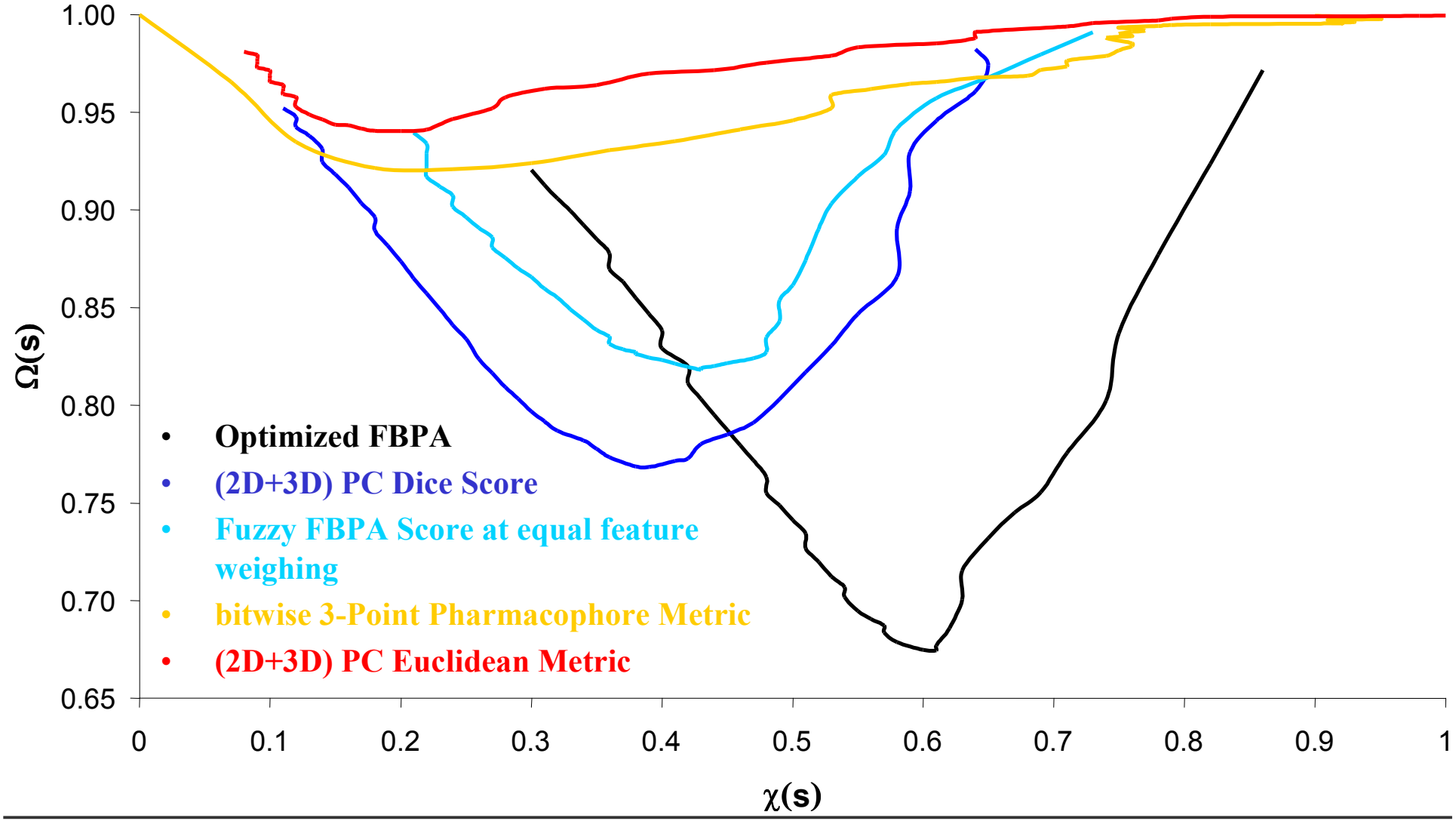
•Plot of $\Omega(s)$ against $\chi(s)$:

–Independent of the similarity scale

–Deeper the curve, higher consistency and then neighborhood behavior

Horvath, D.; Jeandenans, C. *J. Chem.Inf. Comp. Sci.*, 43(2), 2003, p 680-690
Horvath, D.; Jeandenans, C. *J. Chem.Inf. Comp. Sci.*, 43(2), 2003, p 691-698

# Comparison of metrics




Cerep

# Virtual Library: the Potential

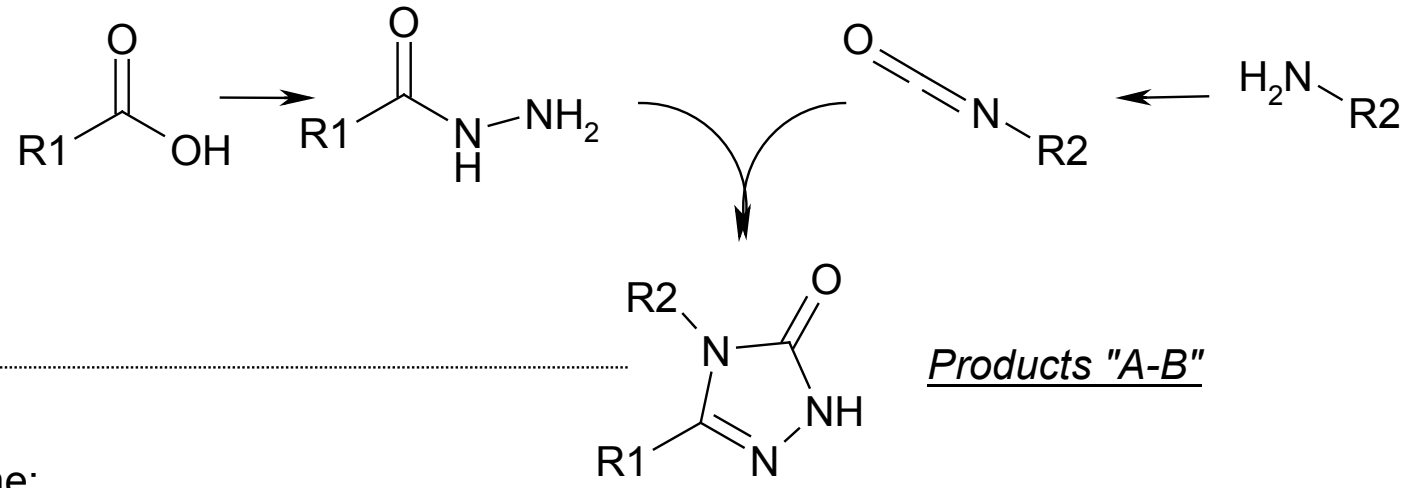
- 100 millions compounds ($10^8$)
  - Mostly combinatorial compounds
  - Chemically feasible
- 3D descriptors
  - FBPA (fuzzy bipolar pharmacophoric autocorrelograms)

# Cerep Library Generation

- Register each building blocks
  - Reactivity assess automatically
  - 3D conformation generation on each building blocks
  - Optimization of the formed bond for the final product
- 3D descriptors
  - Calculation very fast
  - Descriptors stored

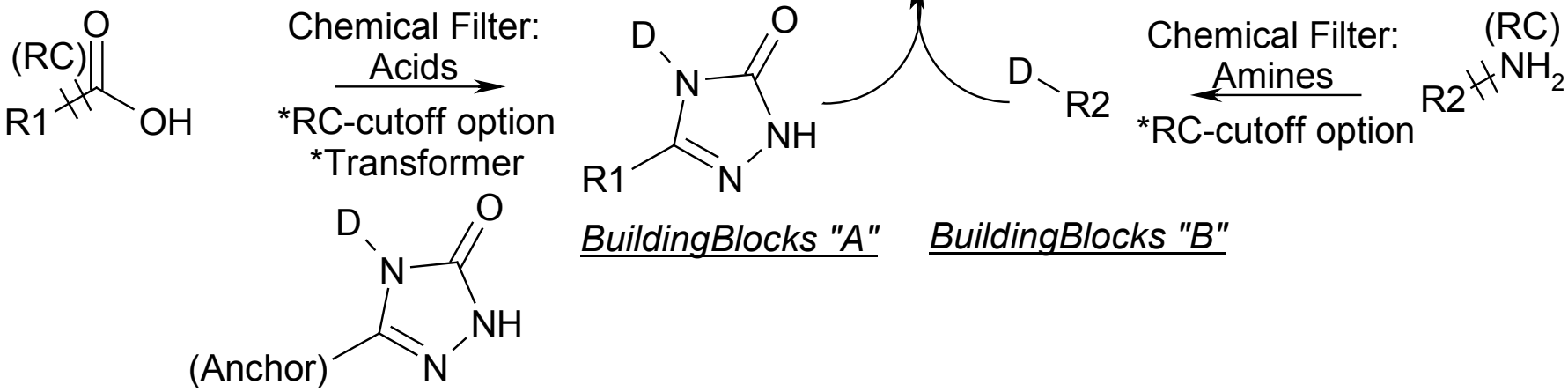
# Library Generation

Reaction Scheme:




*Products "A-B"*

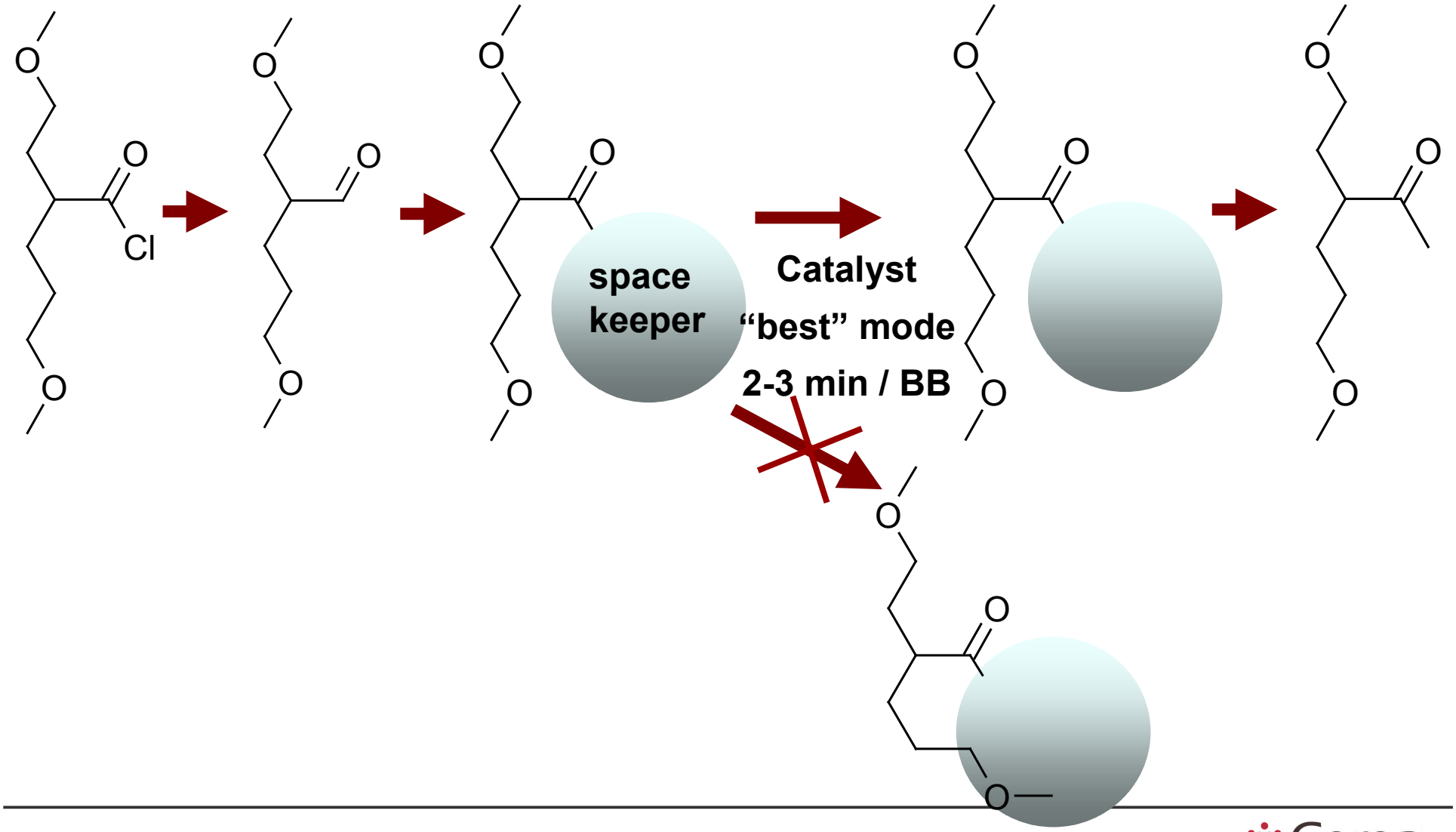
Formal Buildup Scheme:



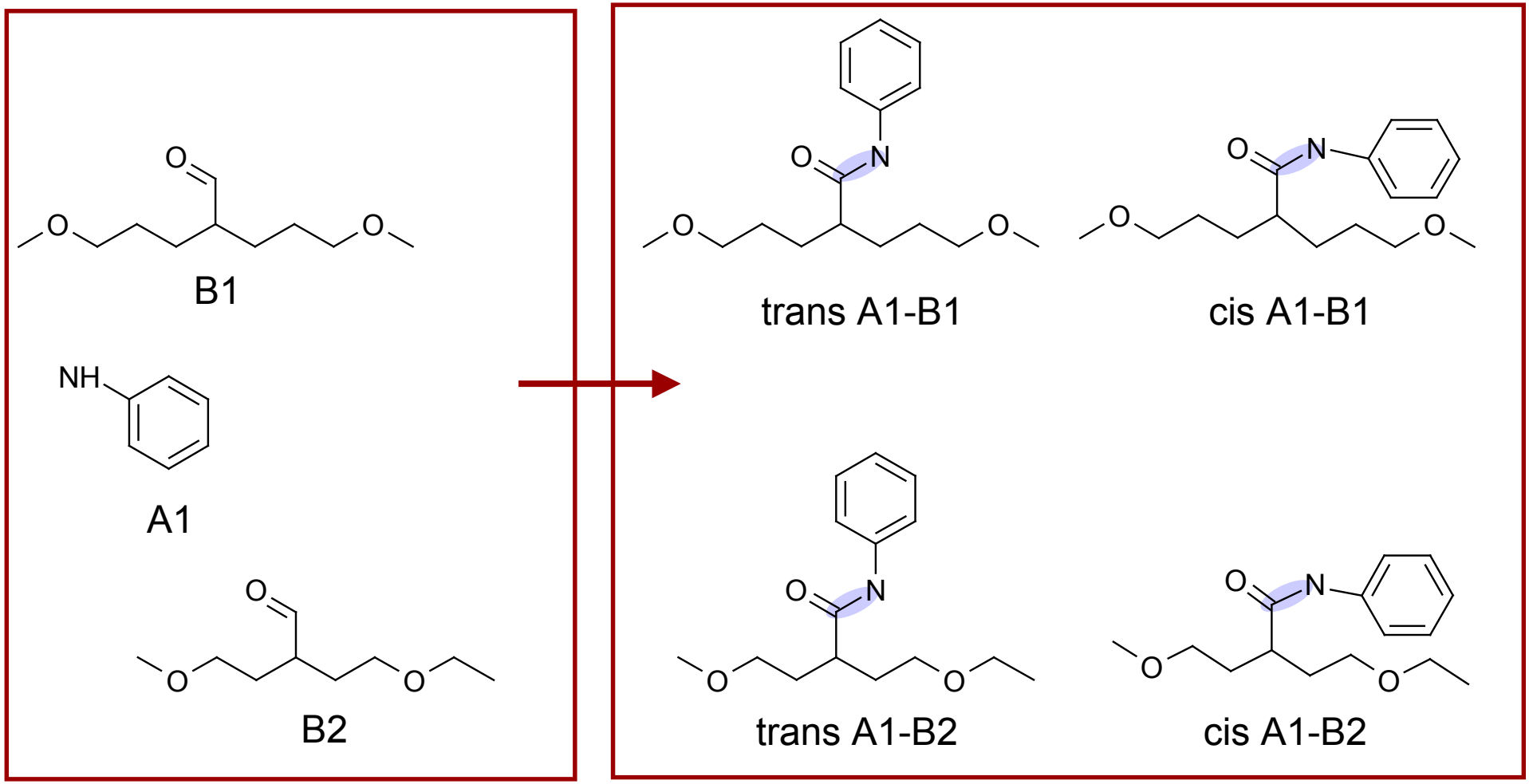

*BuildingBlocks "A"* *BuildingBlocks "B"*

Cerep

# Cerep Library Generation



# Ghost Database Mechanism



# Compounds in the Potential Library

- More than 5000 Monomers
  - In House: multi-step, original, focused, …
  - Commercial: classic and exotic providers
  - Amines, Acides, Aldehydes, Alcool, Halogenures, …
- More than 10 connecting chemistries, multi-step
  - Amine + Electrophile: amide, sulfonamide, …
  - Reductive Amination
  - Heterocycle synthesis: oxadiazole, thiotriazole
  - …

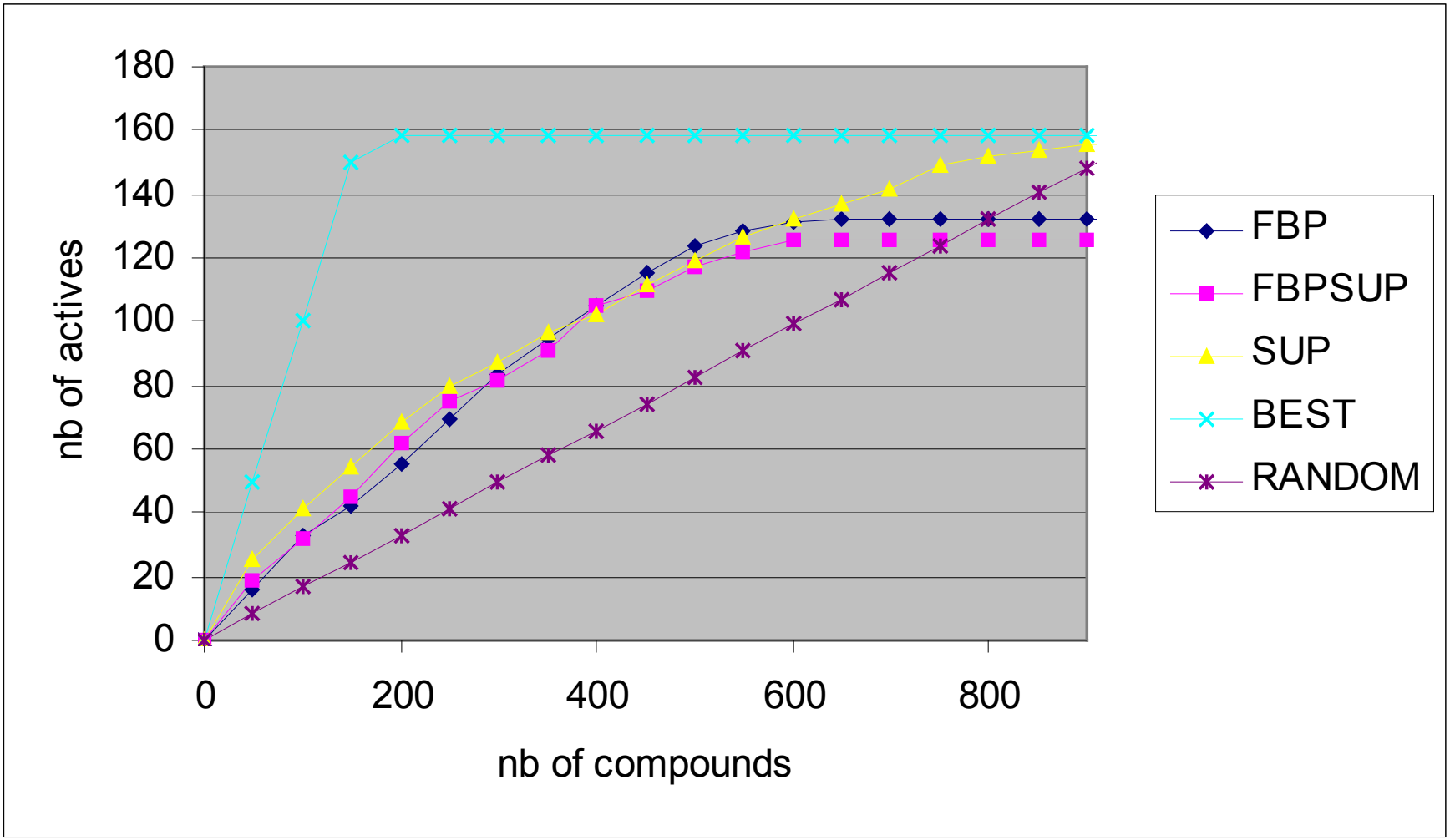
# Virtual Screening at Cerep

- Potential Library
  - Virtual screening really fast
  - Compounds rapidly and easily available
- Quality ?
  - FBPA+Superposition screening compared to superposition
  - Hits retrieved

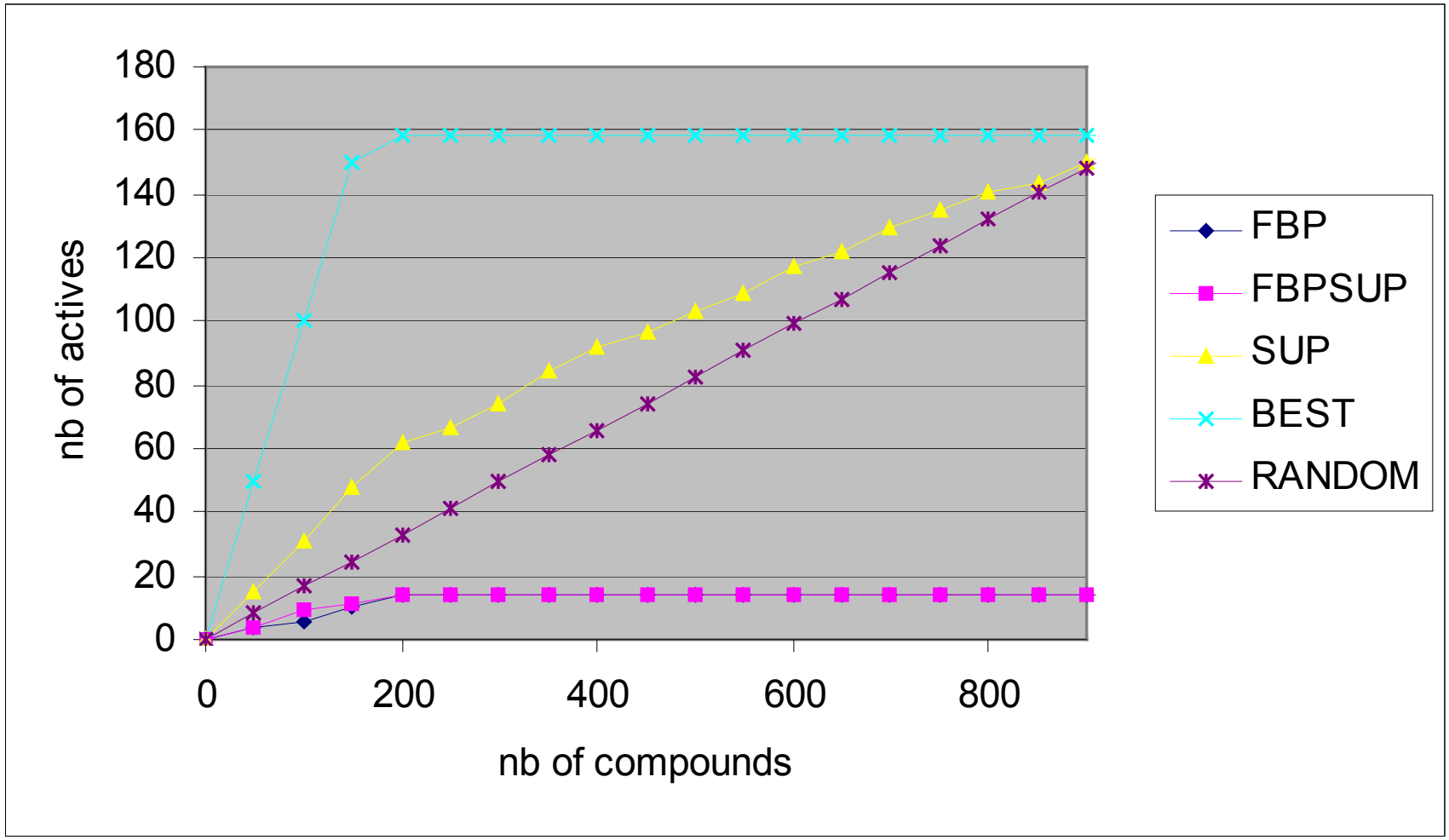
# Example

- 13 GPCR
  - 5HT1a, 5HT2a, 5HT2c, 5HT7, alpha1a, alpha2a, D1, D2, H1, M1, M2, M3, mu
- Reference coming from BioPrint®
  - Best inhibitor chosen for each target
- Test Library: 960 compounds
  - 240 GPCR focused
  - 720 Diverse library
  - Assayed on 21 GPCR
- Virtual screening
  - FBPA
  - FBPA + Superposition
  - Superposition

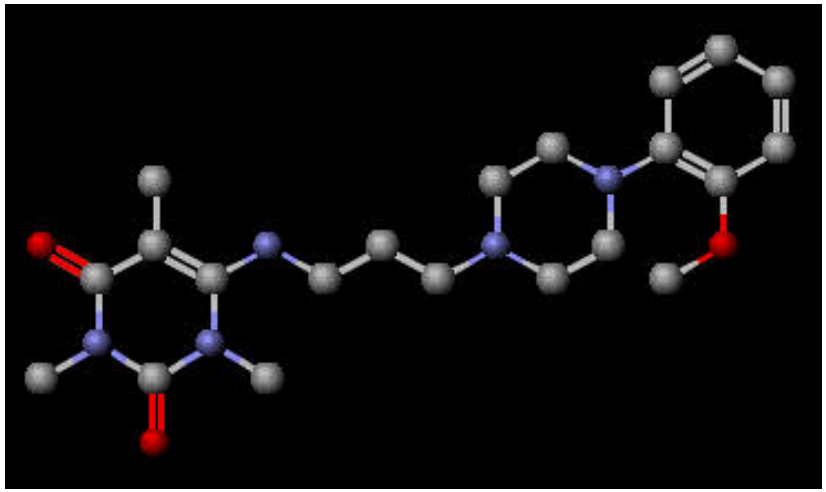
# Results on 5HT1a



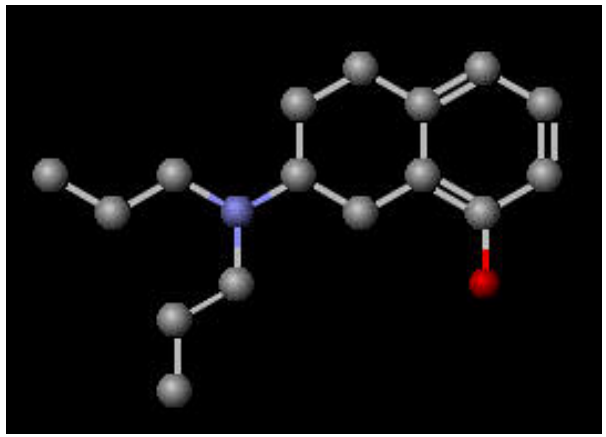
# Results on 5HT1a: Reference 2



# Why so bad Behavior? Choice of the Reference

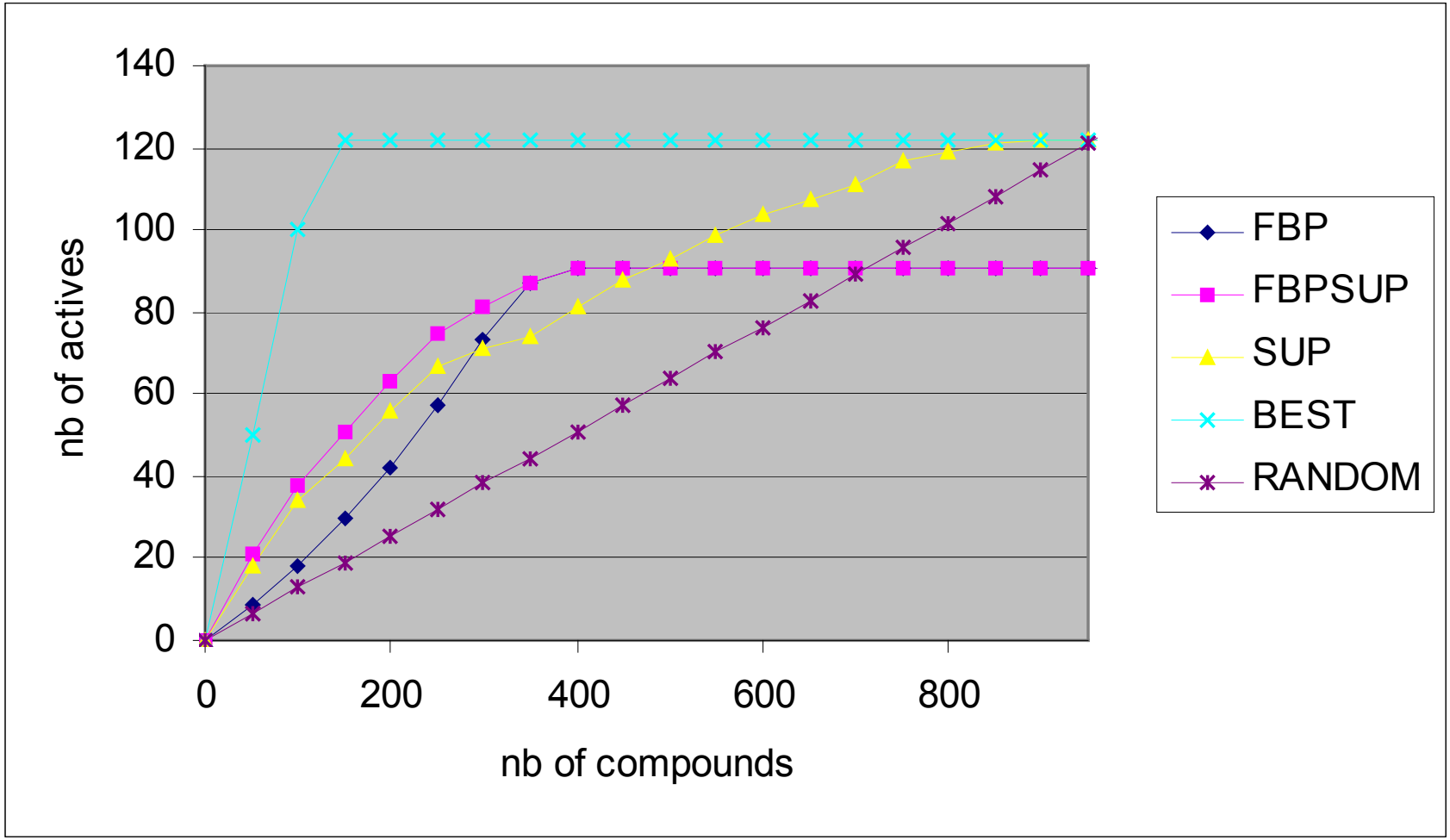


Reference Compound 1

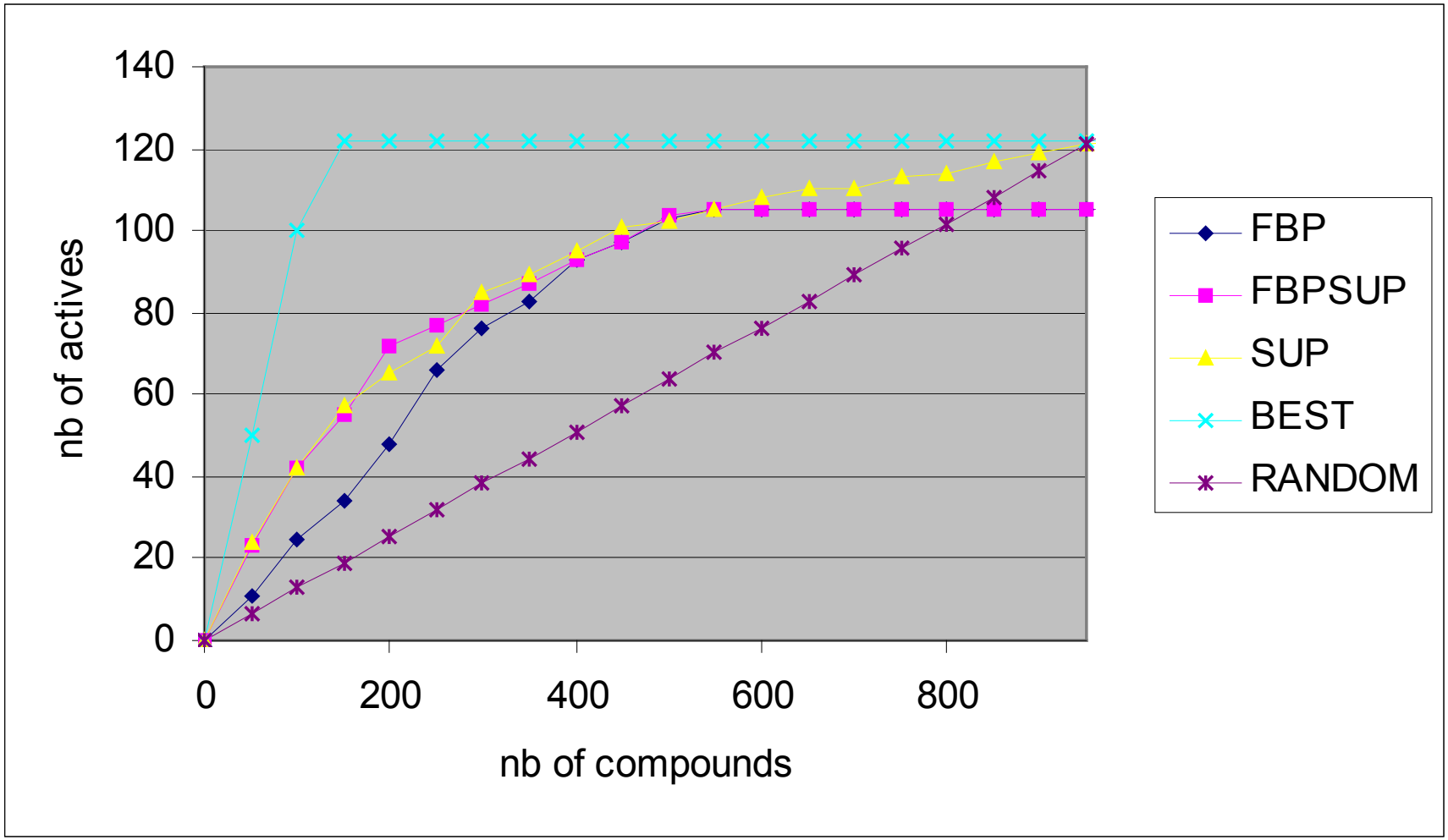


Reference Compound 2

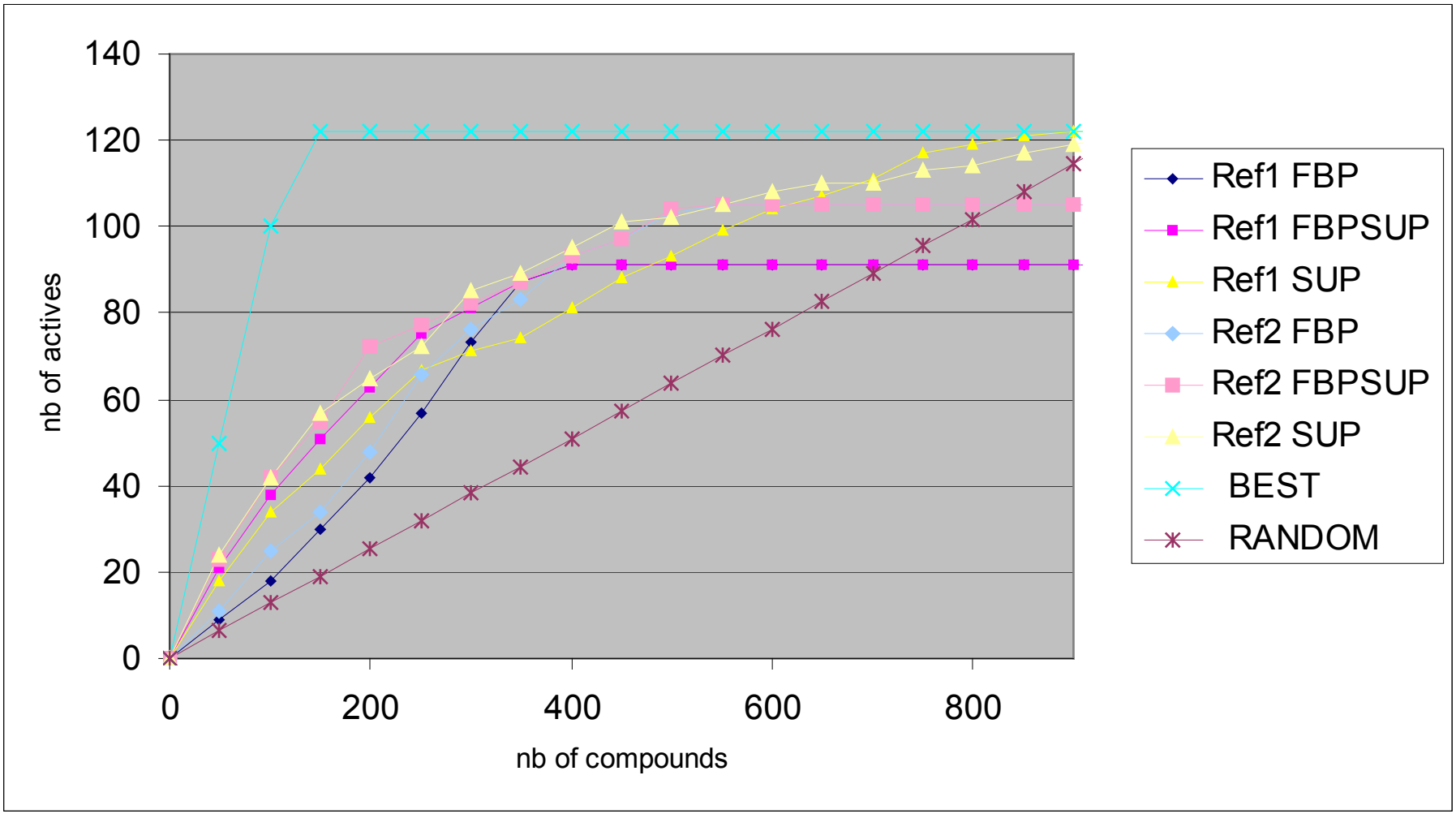
# Results on 5HT2a



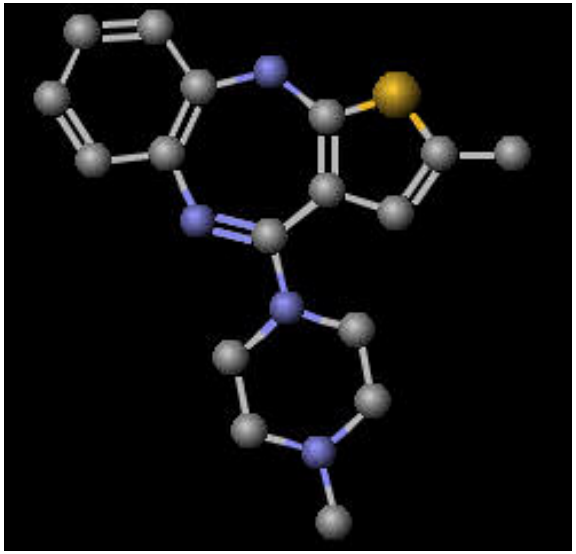
# Results on 5HT2a: Reference 2



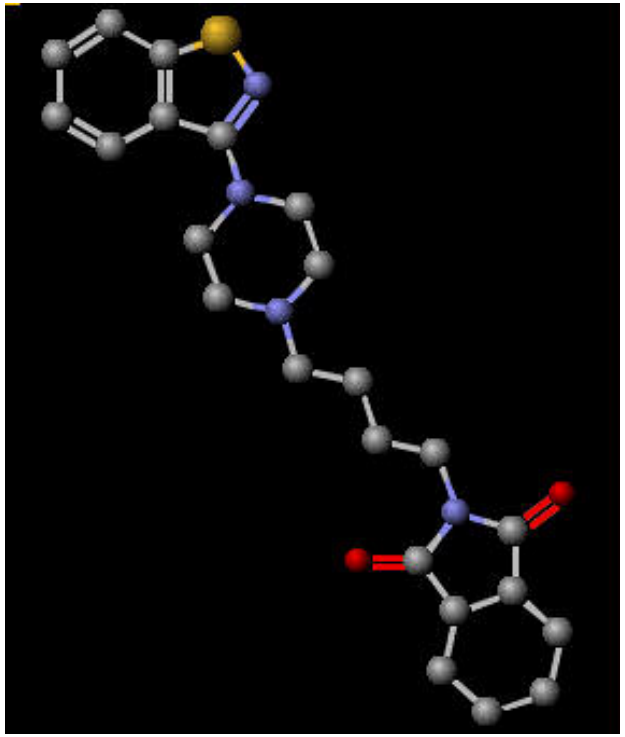
# Results on 5HT2a: comparison



# Choice of the Reference



Reference Compound 1



Reference Compound 2

# Importance of the Reference

- Better to work with several known ligands
- Virtual screening using each ligand
  - Consensus score for the better hits
- Pharmacophore hypothesis
  - Use information known about the active compounds
  - Superposition of the actives
  - Creation of the pharmacophore hypothesis
  - Virtual screening using the pharmacophore hypothesis

# Virtual Screening at Cerep: Conclusion

- Potential Library
  - Large choice of compounds in number and diversity
  - Easy to access the virtual hits
- Ghost Database Mechanism
  - Rapid to generate large virtual combinatorial library
  - Rapid to screen the library
- Quality of the results
  - Good enrichment
  - New chemical families

# End of presentation