

Á. Peragovics^{1,2}, M. Vigh-Smeller¹, Z. Simon^{1,2}, B. Jelinek¹, P. Hári²,
P. Czobor³, Cs. Hetényi¹, A. Málnási-Csizmadia¹

¹Department of Biochemistry, Institute of Biology, Eötvös Loránd University, Pázmány Péter sétány 1/C, H-1117 Budapest, Hungary

²Delta Informatika Inc., Szentendrei út 39-53., H-1033 Budapest, Hungary

³Department of Psychiatry and Psychotherapy, Semmelweis University, Balassa utca 6, H-1083 Budapest, Hungary

Introduction

- Only few systematic effect-specific screening methods exist.
- Potency is strongly correlated with molecular size.
- Ligand Efficiency (LE) is a simple metric for reducing this association.
- Common formulas of LE (assuming direct proportionality):

$$\frac{\Delta G_b}{MW} \quad \frac{\Delta G_b}{N_{HA}} \quad \frac{pK_d}{MW}$$

Problems:

- Residual dependence typically remains
- The potency of the molecules is measured by using different targets → comparison?

Results

Elimination of $\sqrt{\Delta G_b}$ - MW relationship

Hyperbolic fittings were performed and completely size-independent statistical LE values were introduced (LE_{std} and LE_{min}).

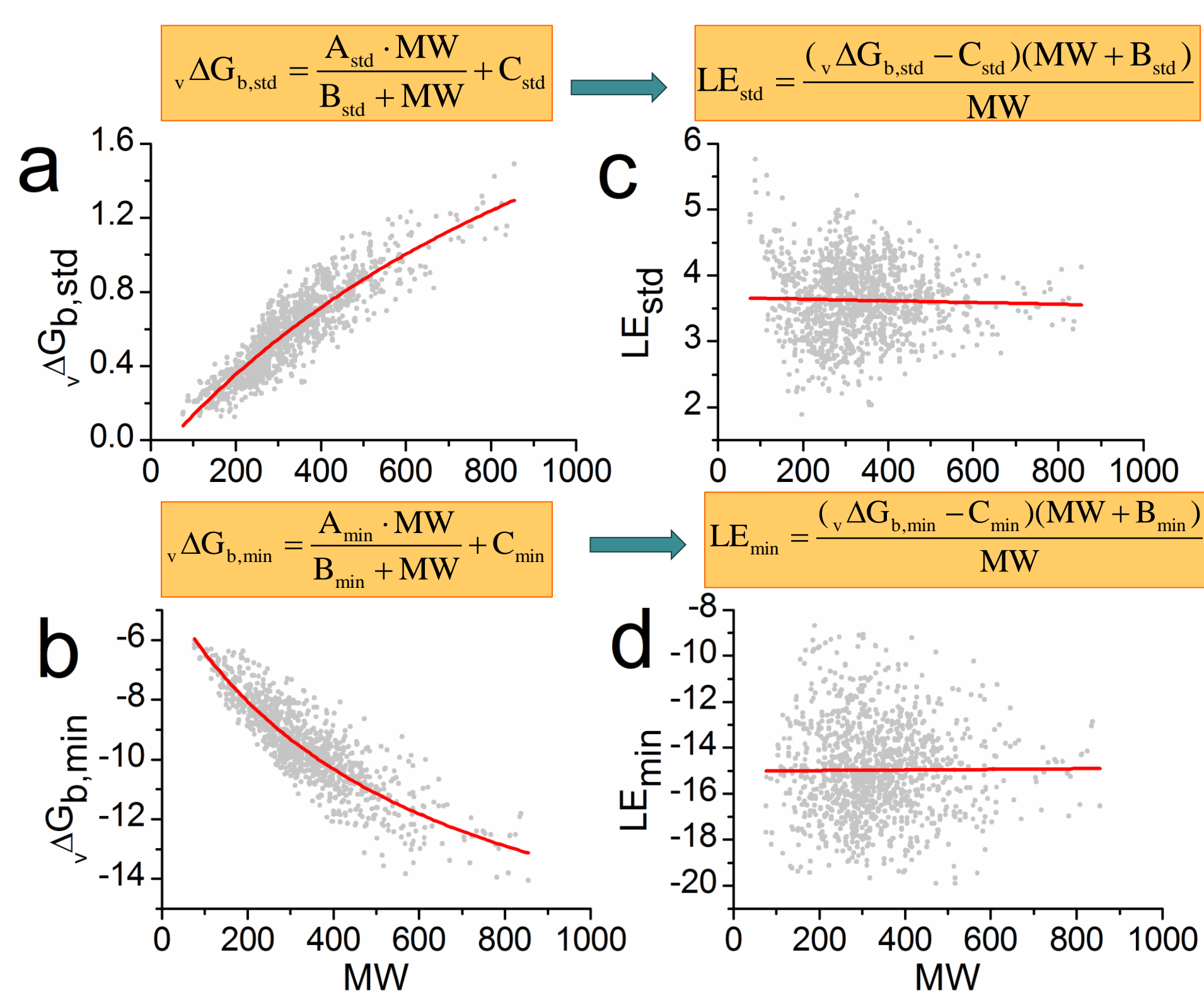


Fig 1. Plots of the $\sqrt{\Delta G_b}$ s and LEs presenting their relations.

Effect-specificity of LE values

In many cases molecules with the same effect tend to have similar statistical LEs thus they form a cohesive group on the LE_{min} - LE_{std} scatterplot.

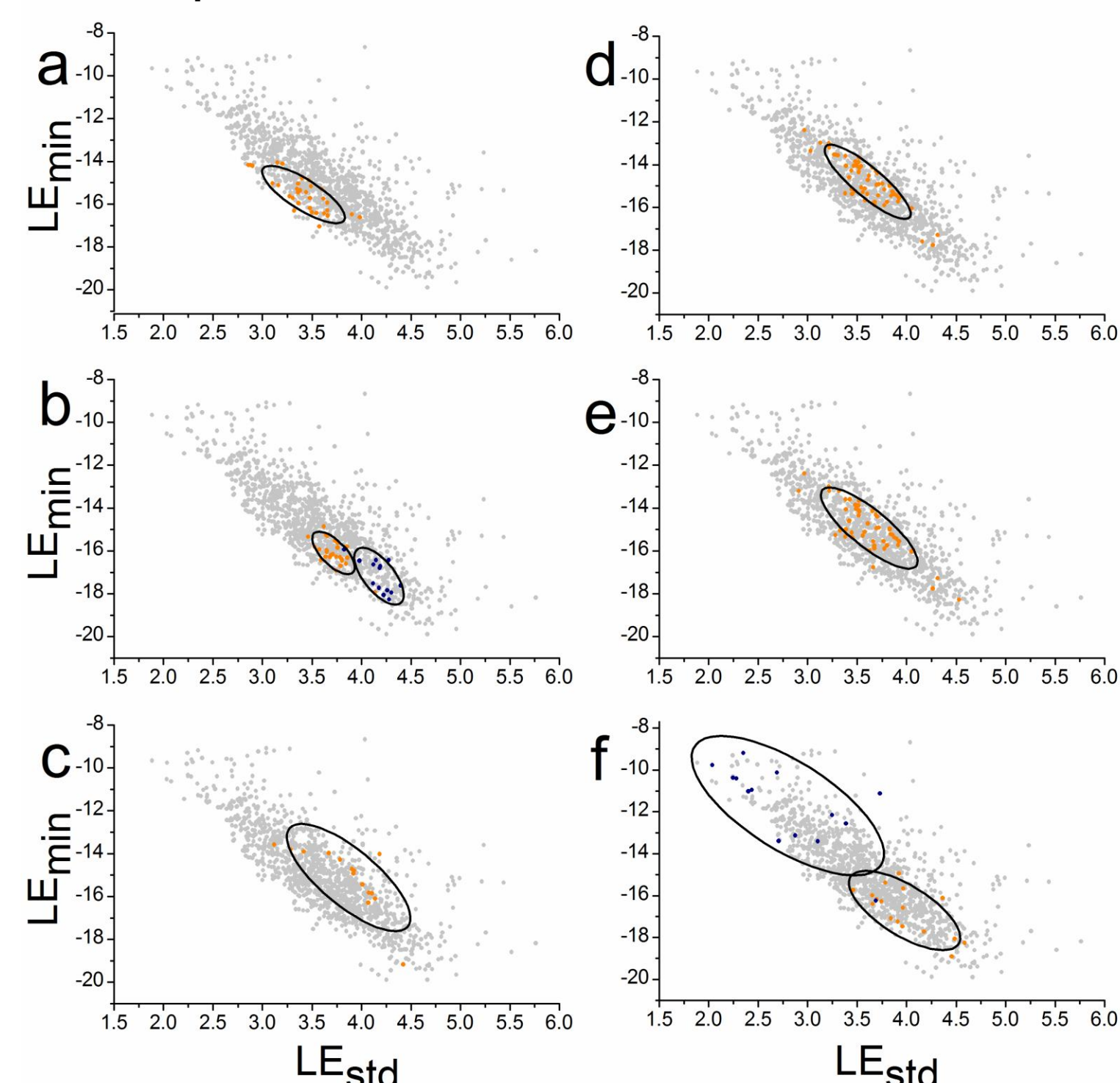


Fig 2. Selected subgroups in the FDA approved drug molecules. 70% confidence ellipses are shown for the data points.

a: benzodiazepines

b: steroidal anti-inflammatory agents (orange) and contraceptives (blue)

c: angiotensin-converting enzyme inhibitors

d: dopamine antagonists

e: antipsychotics

f: alkylating (orange) and hormonal antineoplastic agents (blue)

Aim

- Complete elimination of the residual dependence between LEs and MW:
 - Identify an adequate function describing the ΔG_b -MW relationship
 - Correct ΔG_b with the exact parameters of this function
- To assure comparability, a diverse non-target protein set was used to obtain LEs.
- The binding affinity spectra of the molecules are produced by molecular docking.
- LEs contain general, statistical information on the binding characteristics of the drugs.
- LEs hold out a promise to being effect-specific.
- Estimation of statistical LE values by simple chemical properties

Multilinear regressions

In order to eliminate the extensive docking process, linear models were created to predict LEs by simple chemical descriptors.

General form of the final model:

$$LE_{chem} = \sum_{i=1}^n a_i \cdot p_i + c$$

LE_{chem} : statistical LE value calculated by the chemical properties
 a_i : regression parameter of the i -th chemical property p_i
 c : intercept
 n : number of chemical properties in the model

| Chemical descriptor | $LE_{min,chem}$ | $LE_{std,chem}$ |
|--------------------------|-----------------|-----------------|
| Acceptor count | ✓ | ✓ |
| Aliphatic bond count | ✓ | ✓ |
| Aromatic ring count | ✓ | ✓ |
| Carbo ring count | ✓ | ✓ |
| Log D | ✓ | ✓ |
| Log P | ✓ | ✓ |
| Molecular polarizability | ✓ | ✓ |
| Randic index | ✓ | ✓ |
| Ring atom count | ✓ | ✓ |
| Rotatable bond count | ✓ | ✓ |
| TPSA | ✓ | ✓ |

Table 1. Chemical descriptors of the new linear models.

Step-by-step library filtering method

- Centralization of the data for a selected effect
- Performing PCA (principal component analysis) for the selected effect
- Defining an ellipse which covers at least 70% of the drugs registered to this effect

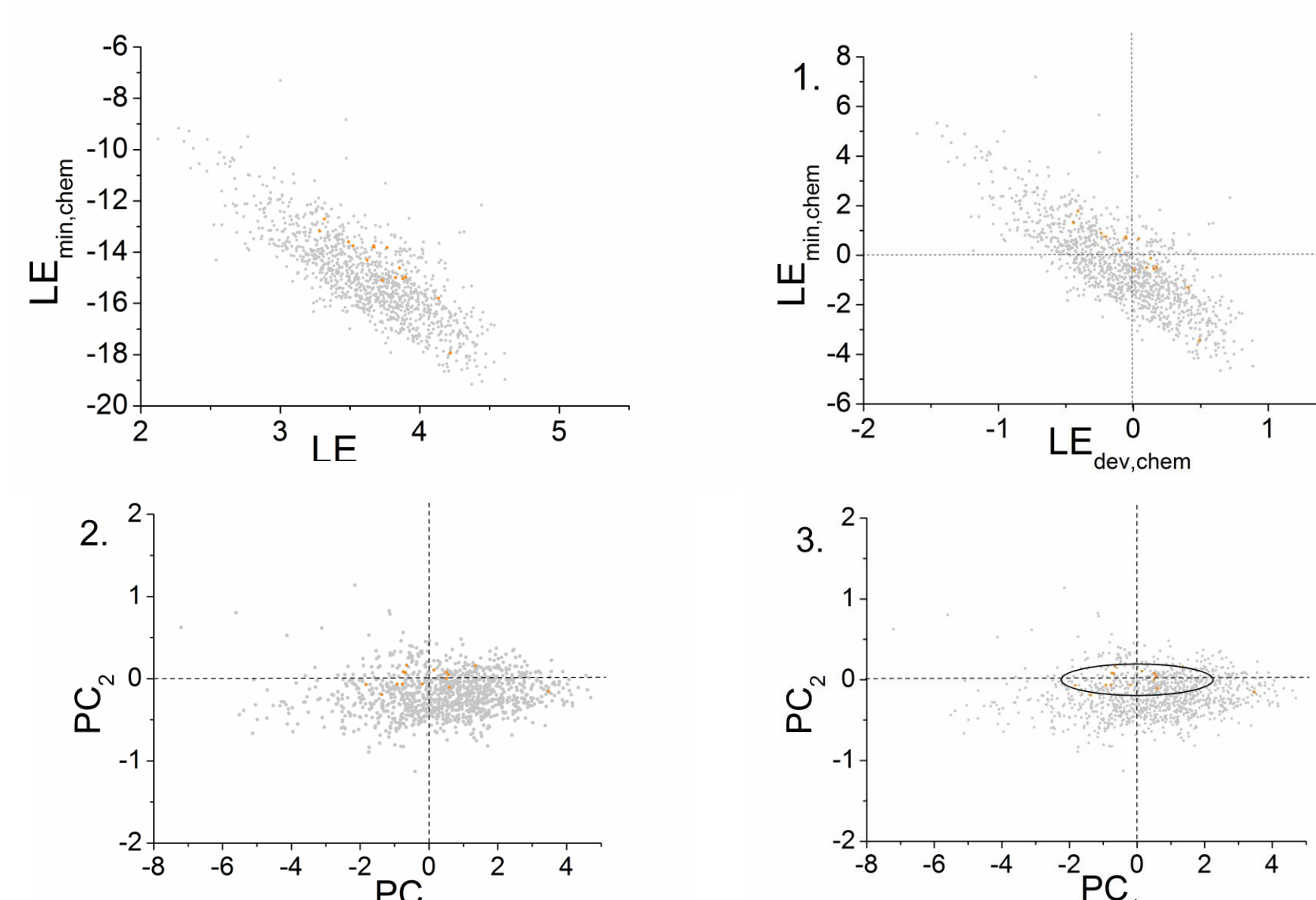


Fig 3. Determination of filtering criteria.

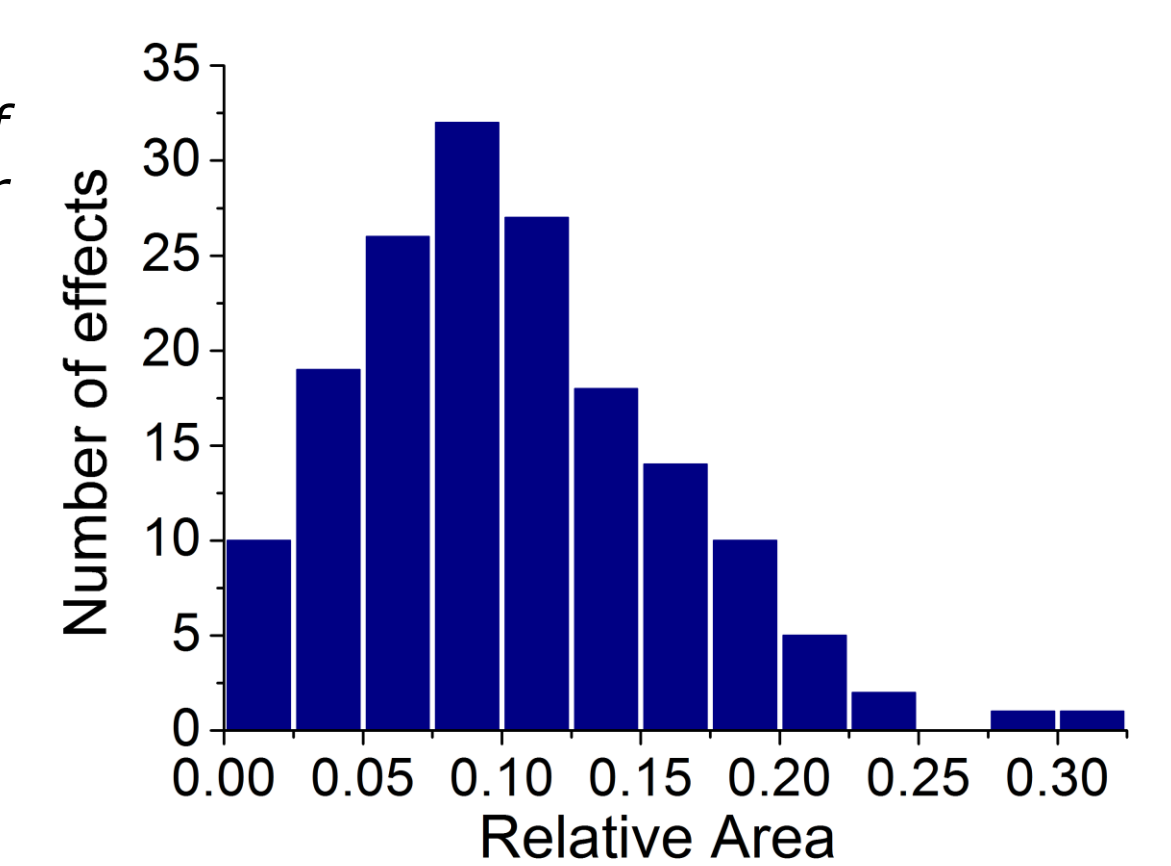
Methods

- Molecular dockings of 1,090 FDA approved drugs to 137 non-target proteins → virtual binding free energy data ($\sqrt{\Delta G_b}$)
- Calculation of the minima and the standard deviations of the $N=137$ $\sqrt{\Delta G_b}$ values for each drug:
 - $\sqrt{\Delta G_{b,min}}$ → maximal binding affinity of a drug
 - $\sqrt{\Delta G_{b,std}}$ → diversity of its affinity spectrum (potency for specificity)
- Extraction of pharmacological information on the 1,090 small molecule drugs from DrugBank database and manual revision → 165 effect category
- Calculation of 41 physicochemical and topological descriptors for each molecule by the Calculator plugins of the ChemAxon Jchem Base software v5.3.7

Enrichment of the used effects

Relative areas were calculated for each effect which contain information about the area they cover compared to the area of the whole drug set.

Fig 4. Histogram of relative area values for the studied 165 effects.



On the average around 9% enrichment can be achieved, however several effects can be screened even more exceedingly.

| Effect category | Relative area |
|---------------------------|---------------|
| Barbiturates | 0.016 |
| Antipsychotics | 0.045 |
| Dopamine Antagonists | 0.049 |
| Serotonin Antagonists | 0.064 |
| Cyclooxygenase Inhibitors | 0.066 |

Table 2. Relative area values of selected effect categories

Conclusion

- Effect-specific filtering of a given chemical space by using statistical LEs
- Estimation of statistical LEs by simple chemical descriptors eliminating the large number of docking runs
- Unique filtering method by itself or second step after the filtering for molecules with appropriate pharmacokinetic properties

Contact Information

Department of Biochemistry Eötvös University

H-1117 Pázmány Péter sétány 1/C Budapest, Hungary

Phone: +36-1-372-2500 ext. 8780

Fax: +36-1-381-2172 E-mail: malna@elte.hu