

정량적 구조-활성 상관 관계와 생리학 기반 약물동태를 사용한 새로운 선도물질 최적화 전략

변진주 · 박민호 · 신석호 · 신영근[#]

충남대학교 약학대학

(Received April 5, 2015; Revised July 22, 2015; Accepted July 31, 2015)

Novel Lead Optimization Strategy Using Quantitative Structure-Activity Relationship and Physiologically-Based Pharmacokinetics Modeling

Jin-Ju Byeon, Min-Ho Park, Seok-Ho Shin and Young Geun Shin[#]
College of Pharmacy, Chungnam National University, Daejeon 305-764, Korea

Abstract — The purpose of this study is to demonstrate how lead compounds are best optimized with the application of *in silico* QSAR and PBPK modeling at the early drug discovery stage. Several predictive QSAR models such as IC₅₀ potency model, intrinsic clearance model and brain penetration model were built and applied to a set of virtually synthesized library of the BACE1 inhibitors. Selected candidate compounds were also applied to the PBPK modeling for comparison between the predicted animal pharmacokinetic parameters and the observed ones *in vivo*. This novel lead optimization strategy using QSAR and PBPK modelings could be helpful to expedite the drug discovery process.

Keywords □ lead optimization, QSAR modeling, PBPK modeling, StarDropTM, ADMET predictorTM, GastroPlusTM

약물의 물리화학적 성질과 Drug Metabolism and Pharmacokinetics(DMPK) 특성은 신약검색과정에서 아주 중요하다.¹⁾ 시판 중인 경구 의약품의 대부분이 목표 단백질에 대한 효능이 클 뿐만 아니라 약물의 분자량, 물리화학적 · DMPK 특성 등이 모두 알맞게 균형을 이루고 있다는 사실이 또한 이를 뒷받침하고 있다.²⁾ 그러나 종래의 신약개발 과정은 단백질에 대한 Ki 값 또는 IC₅₀ 값을 통해 선도물질을 확인한 후 그와 유사한 수천 또는 수만 개의 화합물들을 합성하여 그 물질에 대한 다양한 *in vitro/in vivo* 실험을 진행하여 최적화하는 방법을 사용해 왔다. 이러한 방법은 Absorption Distribution Metabolism Excretion (ADME) 특성과 Pharmacokinetics(PK) 특성이 개선된 후보물질을 찾는 데 있어서 시간이 많이 소요될 뿐만 아니라 이에 비례하여 비용도 크게 늘어난다. 이러한 이유로 주어진 시간과 비용

내에서 신약검색에 대한 효율을 높이기 위해 화합물을 합성하기 전에 물리화학적 성질과 *in vitro* ADME 특성, 그리고 pharmacokinetics(PK) 프로파일을 예측하는 방법이 강구되어 왔고 이러한 방법 중 하나로 컴퓨터 프로그램을 이용하여 위와 같은 특성을 미리 예측하는 *in silico* 모델링이 각광받고 있다.

이러한 *in silico* 모델링은 최근 10~15년 동안 신약검색 및 개발과정에서 약물의 특성 - 물리화학적 성질, 대사, 투과도, 수송체, PK, 안정성 - 을 예측하기 위해 널리 사용되어 왔다.³⁾ *In silico* 모델링은 단기간 내에 수천-수만 개의 화합물을 가상으로 합성하여 그 중 목표로 하는 ADME/PK 특성을 기준으로 순위를 부여할 수 있고, 이 중 상위권에 있는 화합물들을 먼저 선택적으로 합성한다면 보다 효율적으로 시간과 비용을 절감하면서 후보물질을 도출해낼 수 있다. 나아가 *in vitro/in vivo* assay 자료를 효율적으로 활용하여 여러 동물 종뿐 아니라 사람의 약동학적 프로파일도 예측할 수 있다. 이와 관련된 *in silico* 모델링으로 물리화학적 특성을 분자 구조 기반 함수로서 해석하는 Quantitative Structure-Activity Relationships(QSAR) 모델링과 동물의 생리화학적 특성을 바탕으로 약동학적 프로파일을 예측하는 Physiologically-based pharmacokinetics(PBPK) 모델링은

[#]Corresponding Author

Young Geun Shin

College of Pharmacy, Chungnam National University, Daejeon 305-764, Korea

Tel.: 042-821-5931 Fax.: 042-823-6566

E-mail: yshin@cnu.ac.kr

선도물질을 최적화 하는 데 크게 기여 할 수 있다. 특히 PBPK 모델링은 최근에 US FDA 가이드라인⁴⁾ 등에서 임상개발 약물의 약물-약물 상호작용을 초기에 예측하는 데에 있어서도 널리 사용하도록 권장되고 있다.⁵⁾ 이러한 이유로 다국적 제약사에서는 보다 효율적으로 선도물질을 최적화하고 임상에서의 응용을 넓히는 목적으로 *in silico* 모델링을 적극 활용하는 추세이다.^{3,6-8)}

본 연구에서는 Alzheimer's disease(AD) 치료 전략의 한 표적 단백질인 β -site APP-cleaving enzyme 1(BACE1)의 inhibitor⁹⁾로 알려진 물질들을 가지고 본 연구의 방법론에 적용하였다. 신경성 퇴행 질환 중 가장 많이 알려진 AD의 주요 특징은 amyloid- β (A β) peptide⁹⁾가 뇌 내에 비가역적으로 축적이 되는 것으로 알려져 있고, 따라서 이의 생성을 줄이기 위해 BACE1에 대한 inhibitor를 찾아서 AD의 치료를 위한 선도물질들과 최적화된 후보물질들을 제안하였다. 본 논문에서는 새로운 선도물질 최적화 전략의 Proof of concept을 구현하기 위해 특정 논문⁹⁾에 제시된 BACE1 억제물질들을 대상으로 *in vitro* efficacy 값을 이용하여 다양한 QSAR 모델링을 구축하는 전략을 소개하고, 아울러 논문⁹⁾ 상에서 *in vivo* PK가 진행된 물질을 대상으로 PBPK 모델링을 구현하고자 하였다. 이러한 과정을 통해 다국적 제약사에서 널리 활용하고 있는 *in silico* 기반의 고효율 선도물질 최적화 전략을 소개 하고자 한다.

실험방법

QSAR 모델링을 통한 선도물질 최적화

QSAR(Quantitative Structure-Activity Relationships)란 정량적 구조-활성 상관관계라고 한다. 이러한 관계를 바탕으로 선도 화합물과 그에 대한 유도체에 QSAR를 적용하면 화합물의 분자 구조가 가지는 물리화학적 특성에 따른 생물학적 활성의 변화를 정량적으로 예측 가능하다.¹⁰⁾ 본 연구에서는 QSAR 모델링을 위해 StarDropTM ver 5.5(Optibrium, 영국)을 사용하였다. 효능 예측 QSAR 모델링을 위해서는 논문⁹⁾에 제시된 특정 scaffold를 가진 57개 BACE1 억제 화합물 구조와 실측치인 약효 IC₅₀를 기반으로 StarDropTM에 내재된 Auto-modeller 모듈을 사용하여 약물구조를 기반으로 한 local 약효 예측 모델을 구축하였다(Fig. 1).

모델 구축을 위해 training set %, validation set %, test set %은 각각 65%, 20%, 15%로 설정하고, data의 splitting technique는 clustering 방법을 사용하여 설정하였다. 나머지 파라미터 값은 StarDropTM 내의 기존값을 사용하여 모델을 만들었다. 총 10가지의 모델이 생성된 가운데 validation set과 test set의 R² 값이 큰 모델 중에서도 각각의 set에 해당하는 화합물이 직선에 가까운 모델인 GPFixed(Gaussian Processes: Fixed) algorithm을 선택하였다(R²=0.85).

가상의 라이브러리를 구축하기 위해서 BACE1 억제 약물들의

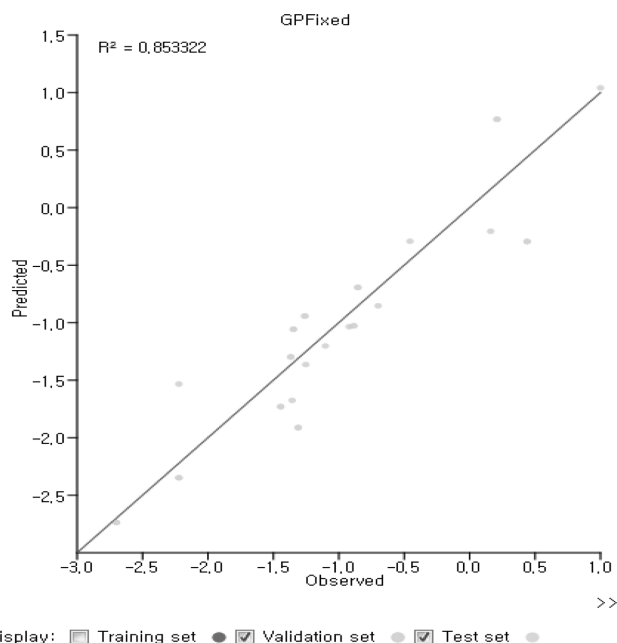


Fig. 1 – Development of a user-defined QSAR model for IC₅₀ prediction using StarDrop Auto-ModellerTM. The best predicted model was produced by the GPFixed algorithm. (R²=0.85; validation set and test set).

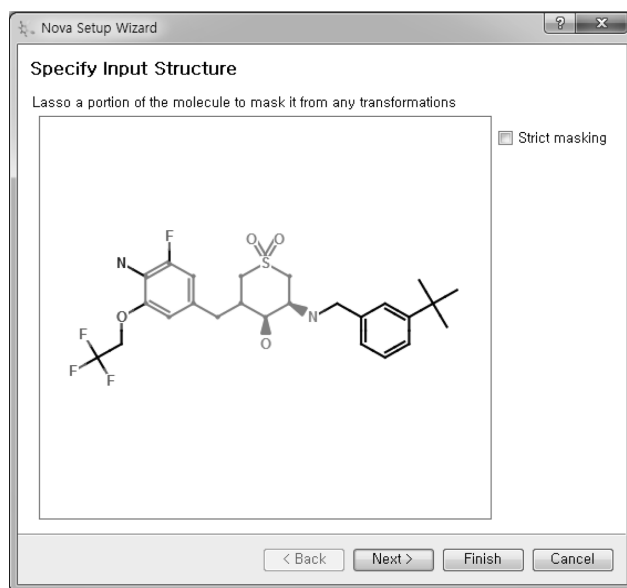


Fig. 2 – *In silico* generation of new library compounds using StarDrop NovaTM.

구조 골격을 먼저 규정한 후, StarDropTM 내의 virtual library synthesis 모듈인 NovaTM 기능을 통해 약 626개의 화합물 구조를 합성하였다(Fig. 2). 그런 다음 가상의 합성 화합물을 대상으로 약효예측모델(log IC₅₀)을 적용하여 가상의 약효를 예측하였다. 가상의 라이브러리 화합물을 합성할 때 총 5번의 단계를 통해서 합성을 하되, 각 단계에서 예측된 logIC₅₀ 값을 기반으로

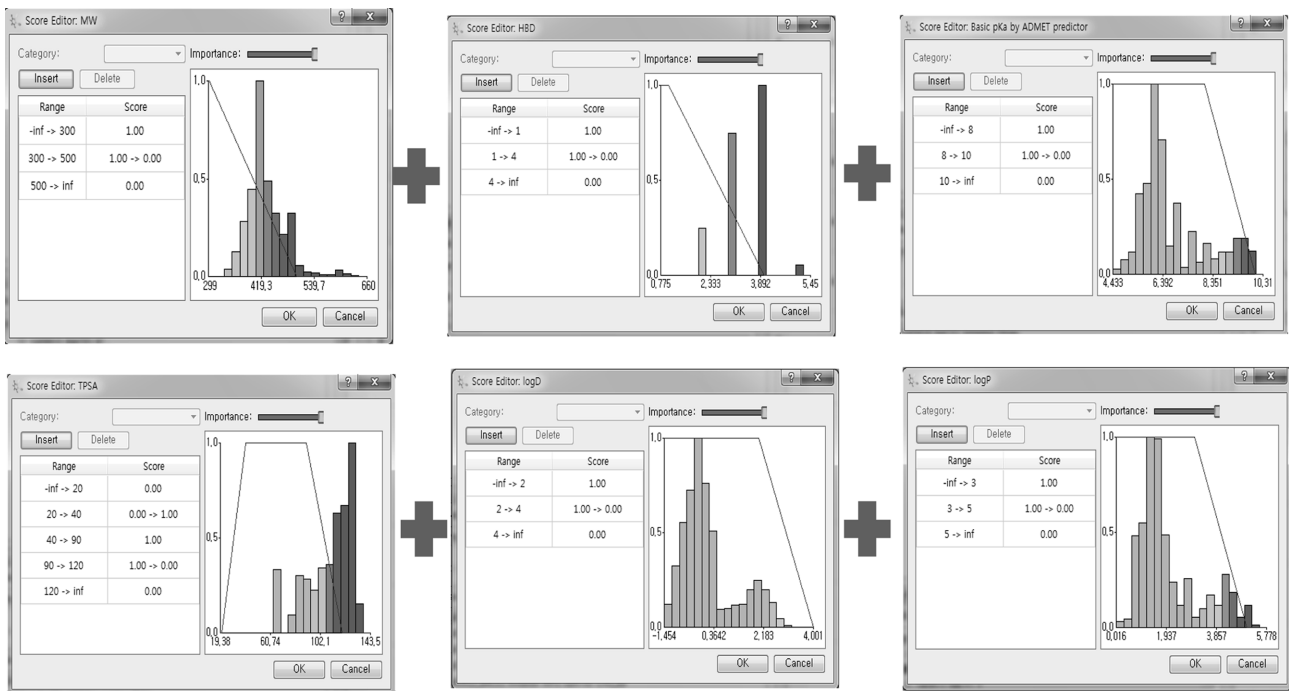


Fig. 3 – Production process of CNS Multi-Parameter Optimization (MPO) score.

100개의 화합물이 선택되도록 하였다. 총 합성된 라이브러리 화합물 중 StarDrop™에 내재된 SMARTS pattern Filter 방법을 이용하여 현실적으로 합성이 가능한 626개의 화합물을 골라내었다. 그런 다음 Central Nervous System Multi Parameter Optimization(CNS MPO) score¹¹⁾와 logIC₅₀, 그리고 ADMET predictor™ 통해 예측된 내인성 청소율의 3가지 점수의 조합으로 626개 화합물들의 순위를 결정하게 하였다. 이 중 특히 AD 등 정신신경질환 치료제로서 반드시 갖추어야 할 특성인 뇌 투과성 여부 평가를 위한 CNS MPO score를 Fig. 3에 나타내었고, 그 공식은 다음과 같다.

$$MPO = \sum \text{Score}(\text{clogP} + \text{clogD} + \text{SPA} + \text{MW} + \text{HBD} + \text{pKa})$$

clogP: calculated Partition coefficient

clogD: Distribution coefficient at pH=7.4

MW: molecular weight

TPSA: topological polar surface area

HBD: number of hydrogen bond donors

pKa: acid dissociation constant at logarithmic scale

PBPK 모델링을 통한 *in vivo* PK profile의 예측

PBPK(Physiologically-based pharmacokinetics)는 생리학적 특성을 고려한 약물동태학으로, 약물의 체내 동태를 파악하기 위하여 단순한 수학적인 모델(compartment model)을 이용하는 전통적인 약물동태학과 달리, 약물의 체내 동태를 정확하게 예측

하기 위하여 생리학적 특성을 고려하여 생체를 구성하는 각 조직 및 장기를 혈류와 연결하여 모델링 하는 것을 말한다.¹²⁾ 이렇게 생리학적인 특징과 약물의 특징, 약물과 생체 반응의 특징 등을 modeling에 도입한다면, 다양한 조건에 따른 혈중 농도 및 표적 장기에서 약물 농도를 예측할 수 있다. 본 연구에서는 PBPK 모델링을 위해 Simulations-Plus사(CA, USA)의 GastroPlus™ (ver.8.5) 를 사용하였다.

논문⁹⁾ 상에서 약효와 target에 대한 선택성을 기반으로 제시한 BACE1 억제 약물 3종(14c, 60h, 60j)⁹⁾을 대상으로 GastroPlus™를 이용하여 생쥐에 2 mg/kg 정맥 투여시 PK 프로파일을 예측하였다. 이들 세 화합물에 대해 생리학적 파라미터로 분자량, clogP, pH 7.62에서의 용해도, 투과도, 혈장 unbound fraction, Concentration_{blood}/Concentration_{plasma}(C_{blood}/C_{plasma}) 비율을 예측했고, PK 파라미터로는 AUC 0-t, V_{dss}, T_{1/2}를 예측하였다. 위의 파라미터 뿐만 아니라 시간에 따른 혈중 약물농도 프로파일을 예측하였다. 모델에 대한 정확도 평가와 관련하여 이론적으로는 실측치와 예측치가 일치하는 것이 가장 이상적이거나 일반적으로 다국적 제약사의 신약검색 및 개발과정에서 널리 통용되고 있는 기준인 실측치를 기준으로 2배 오차범위(two-fold error range(0.5~2)) 내에서 모델을 검증 및 평가하였다.^{13,14)}

실험결과 및 고찰

StarDrop™을 이용하여 QSAR 모델을 만들고 가상으로 새롭

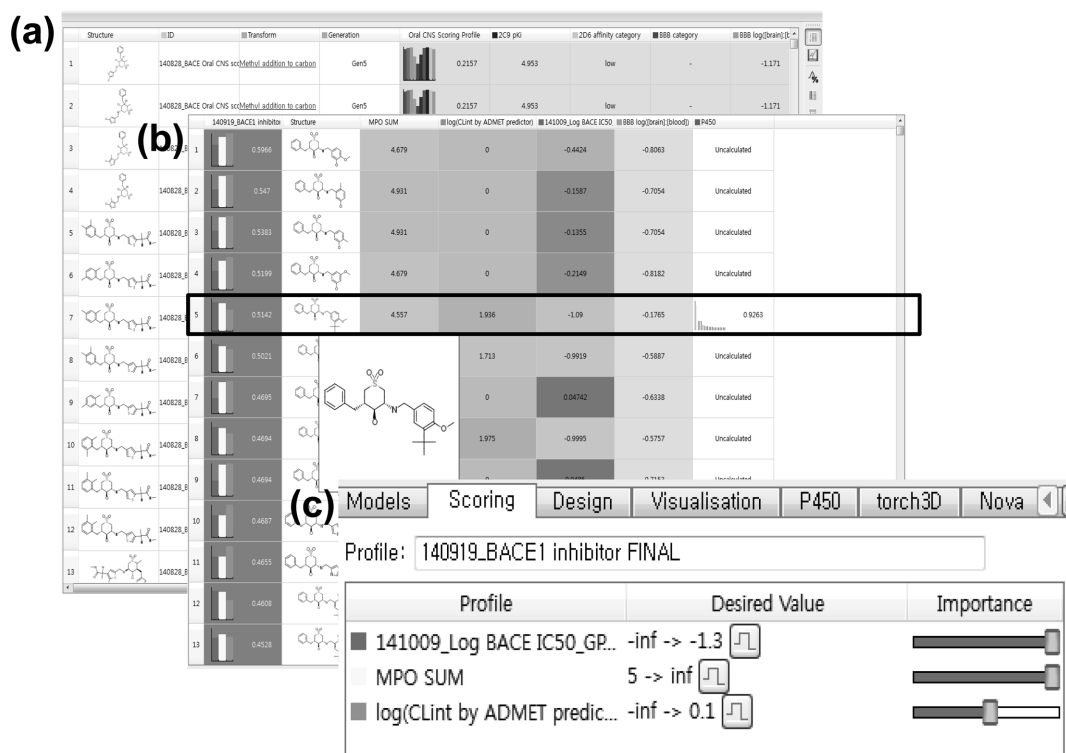


Fig. 4 – (a) 626 compounds virtually generated using StarDrop Nova™. (b) Compounds rank-ordered based on the composite scores. (c) The composite scoring rule used for BACE-1 inhibitors.

게 화합물을 합성한 후 CNS MPO score, logIC₅₀, 그리고 내인성 청소율을 기준으로 순위를 매긴 결과를 Fig. 4에 나타내었다. CNS MPO score는 시판되는 CNS 약물 119개와 Pfizer CNS 후보물질 108개를 분석한 결과, 연관된 물리화학적 성질 6가지를 일컫는 것으로 clogP, clogD, MW, TPSA, HBD, pKa로 이루어져 있다.¹⁵⁾ 이러한 6가지 성질은 동일한 비중으로 합해져 0에서 6사이의 값을 나타내며 숫자가 클수록 CNS 약물로서 적합한 성질을 갖는 것이라 할 수 있다. 최적화된 화합물이 AD 치료제로서 CNS에 작용해야 하므로 CNS에 관한 여러 기준을 균형 있게 맞출 수 있는 CNS MPO score를 최종 점수의 기준 중 하나로 사용하였다. 또한 CNS 투과 후 목표 단백질 BACE1에 대한 억제 약효인 logIC₅₀ 역시 최종 점수의 기준의 하나로 고려하였다. 마지막으로 유효약물 농도의 적절한 유지 및 의미 있는 약동학적인 파라미터를 확보하기 위해 내인성 청소율도 고려하였다.

최종적으로 순위가 매겨진 화합물들의 점수를 Fig. 5와 같이 나타내었다. 가장 점수가 높은 화합물은 왼쪽에 위치했고 이 화합물의 아래쪽 error bar가 겹치는 위치의 화합물¹⁰⁾까지 약 100여개의 화합물이 선택되었다. 이 중에서도 global ADME 모델과 global Brain Blood Barrier(BBB) 투과 모델에 기반하여 후보물질로서 가능성이 있는 화합물을 선택하였다. 논문⁹⁾ 상에서는 효능과 target에 대한 선택성을 기반으로 세 개의 화합물(14c, 60h, 60i)을 선정하여 *in vivo* PK를 진행하였다. 그러나 본 연구에서

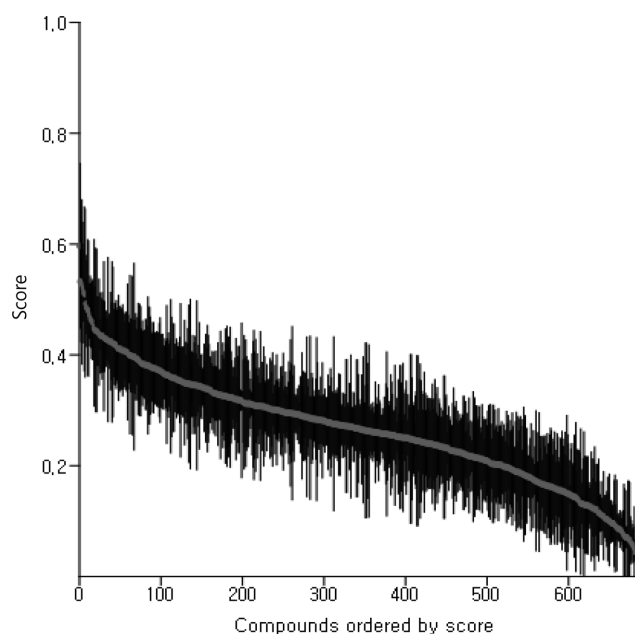


Fig. 5 – Score distribution of all compounds tested by user-defined scoring rule.

최적화된 파라미터 상에서 세 개의 화합물(14c, 60h, 60i)은 각각 210, 420, 625의 순위를 차지하였다. 이는 본 연구에서 제시한 최적화 전략을 사용하였다면 약효, BBB에 대한 투과성, 그리

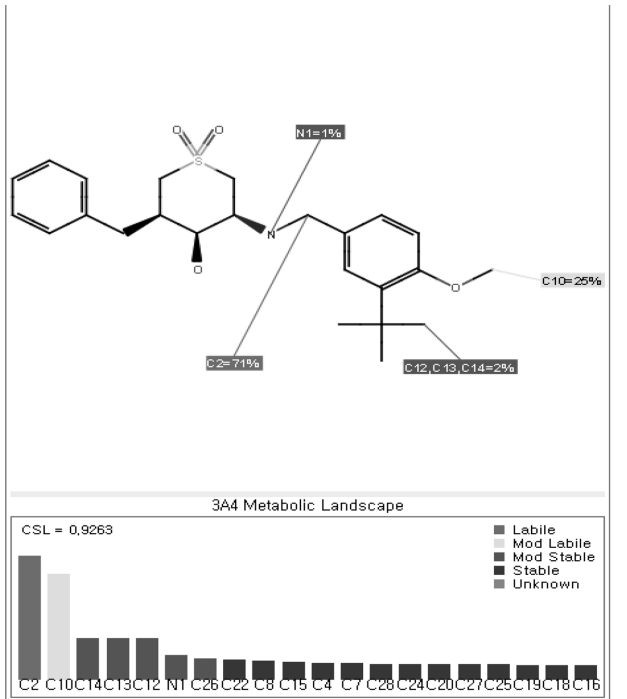


Fig. 6 – Metabolic soft spot analysis of a selected compound using StarDrop P450™.

고 내인성 청소율에 대한 부분을 모두 고려하여 BACE1 저해제로서 조금 더 가능성 있는 물질이 나올 수 있었을 것임을 시사

Table I – Predicted compound specific parameter for 3 candidate compounds (14c, 60h, 60j)

Parameters	14c	60h	60j
MW	495.49	532.60	600.60
Lipophilicity (cLogP)	3.45	4.20	5.63
Solubility(mg/ml @ pH 7.62)	0.01	$4.5 \cdot 10^{-3}$	$9.75 \cdot 10^{-4}$
Permeability (P_{eff} ; $cm/s \cdot 10^4$)	0.69	2.68	3.56
Plasma unbound fraction (f_{up})	11%	17%	18%
C_{blood}/C_{plasma} ratio	0.76	0.71	0.69

한다. Fig. 6은 선택된 특정 화합물 중 약물대사효소(예를 들면 cytochrome P450s 2C9, 2D6, 3A4 등)에 의해 잘 대사되는 취약 부위가 어디인지를 예측한 예이다. 신약 검색에 있어서 대사 안정성을 예측하고, 대사 안정성이 떨어질 경우 어느 부위가 대사를 잘 받는 부위인지를 빨리 알아내는 것은 신약 검색과정에서 매우 중요한 일이며, 구조 기반으로 cytochrome P450s에 의한 주요 대사 부위를 미리 예측하는 것은 최적화 과정에서 크게 도움이 된다. 이러한 일련의 과정을 통해 목표하는 단백질에 대한 효능을 유지하면서 CNS 투과와 청소율이 개선된 화합물의 발견이 가능하다.

GastroPlus™를 이용해 예측된 해부·생리학적 파라미터와 *in vivo* 청소율 값을 고려하여 예측된 물질의 특성과 PK 프로파일은 Table I과 Fig. 7에 나타내었다. Table I에서는 화합물 14c, 60h, 60j 각각에 대해 예측된 분자량, clogP, pH 7.62에서의 용해도, 투과도, 혈장 unbound fraction, C_{blood}/C_{plasma} 비율을 제시하

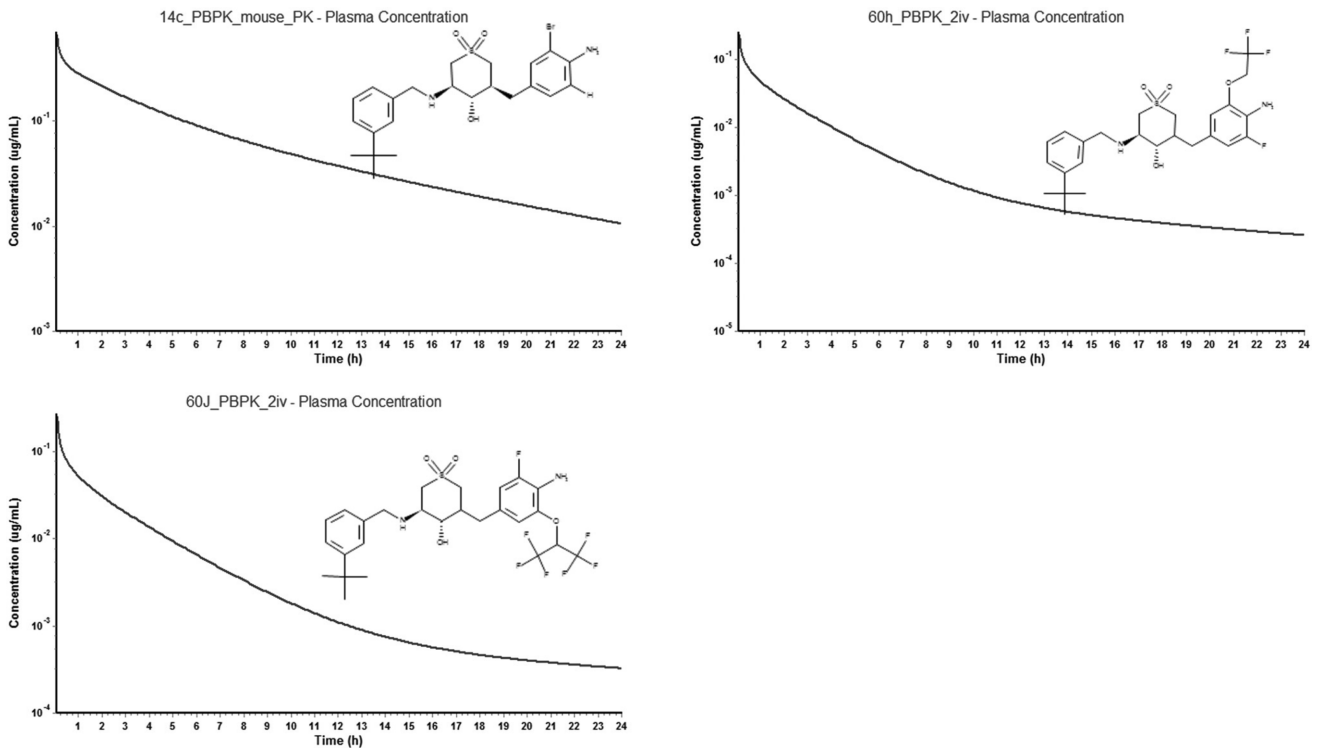


Fig. 7 – Predictive PK profile for 3 candidate compounds (14c, 60h, 60j).

Table II – PBPK modeling result for 3 candidate compounds (14c, 60h, 60j)

Result	14c		60h		60j	
	Observed	Simulated	Observed	Simulated	Observed	Simulated
AUC 0-t ($\mu\text{g} \cdot \text{h}/\text{mL}$)	2.61	1.88	0.20	0.29	0.31	0.33
Vdss (L/kg)	1.60	2.76	7.20	6.40	16.30	9.52
$T_{1/2}$ (hr)	4.00	4.79	1.40	1.55	2.00	2.36

였다. Fig. 7은 최적화 과정을 통해서 도출된 3종의 화합물 14c, 60h, 60j의 PK 프로파일 예측 결과를 혈중농도-시간 그래프로 나타낸 것으로 로그 스케일로 표현하였다.

Table II은 화합물 14c, 60h, 60j에 대해 관찰된 PK 파라미터와 실측치인 PK파라미터를 비교하여 나타내었다. 화합물 14c의 경우 관찰된 AUC 0-t($\mu\text{g} \cdot \text{h}/\text{mL}$)가 2.61, Vdss(L/kg)가 1.60, $T_{1/2}$ (hr)가 4.00이었고 예측된 AUC 0-t($\mu\text{g} \cdot \text{h}/\text{mL}$)가 1.88, Vdss(L/kg)가 2.76, $T_{1/2}$ (hr)가 4.79로 관찰된 PK 파라미터 값, 즉 실측치를 기준으로 AUC 예측치가 2배 오차범위 내에 있었다. 화합물 60h의 경우 관찰된 AUC 0-t($\mu\text{g} \cdot \text{h}/\text{mL}$)가 0.20, Vdss(L/kg)가 7.20, $T_{1/2}$ (hr)가 1.40이었고 예측된 AUC 0-t($\mu\text{g} \cdot \text{h}/\text{mL}$)가 0.29, Vdss(L/kg)가 6.40, $T_{1/2}$ (hr)가 1.55로 이 화합물 역시 실측치를 기준으로 예측치가 2배 오차범위 내에 있었다. 마지막으로 화합물 60j의 경우 관찰된 AUC 0-t($\mu\text{g} \cdot \text{h}/\text{mL}$)가 0.31, Vdss(L/kg)가 16.30, $T_{1/2}$ (hr)가 2.00이었고 예측된 AUC 0-t($\mu\text{g} \cdot \text{h}/\text{mL}$)가 0.33, Vdss(L/kg)가 9.52, $T_{1/2}$ (hr)가 2.36으로 이 역시 실측치를 기준으로 예측치가 2배 오차범위 내에 있었다. 위와 같은 결

과는 적절한 *in silico* 및 *in vitro* data를 토대로 구축한 QSAR 모델링과 PBPK 모델링이 *in vivo* preclinical PK를 허용되는 오차 범위 내에서 적절히 예측 가능하다는 것을 보여주며, 이러한 과정은 많은 노력과 시간이 필요한 후보물질 최적화 시간을 크게 단축할 수 있다는 것을 시사한다.

결론

본 연구는 QSAR 모델링과 PBPK 모델링에 대한 전반적인 소개와 이를 통해 신약검색과정에 대한 새로운 응용 사례를 소개하였다. 다국적 제약사뿐 아니라 특히 최근에는 미국 FDA, 유럽 EMA, 일본 MHLW에서 신약 승인 과정에서 PBPK 모델링 및 시뮬레이션 결과를 약물 상호작용 예측에 활용하고 있고, 일부 자료들은 최종 package insert에도 포함 되는 등¹⁶⁾ 의약품 인허가 당국에서도 많은 관심을 가지고 있다. 본 논문에서는 Fig. 8에 제시된 바와 같이 혁신적 선도물질 최적화 전략에서 QSAR/PBPK *in silico* 모델링이 신약 개발 초기 단계에서 후보물질을

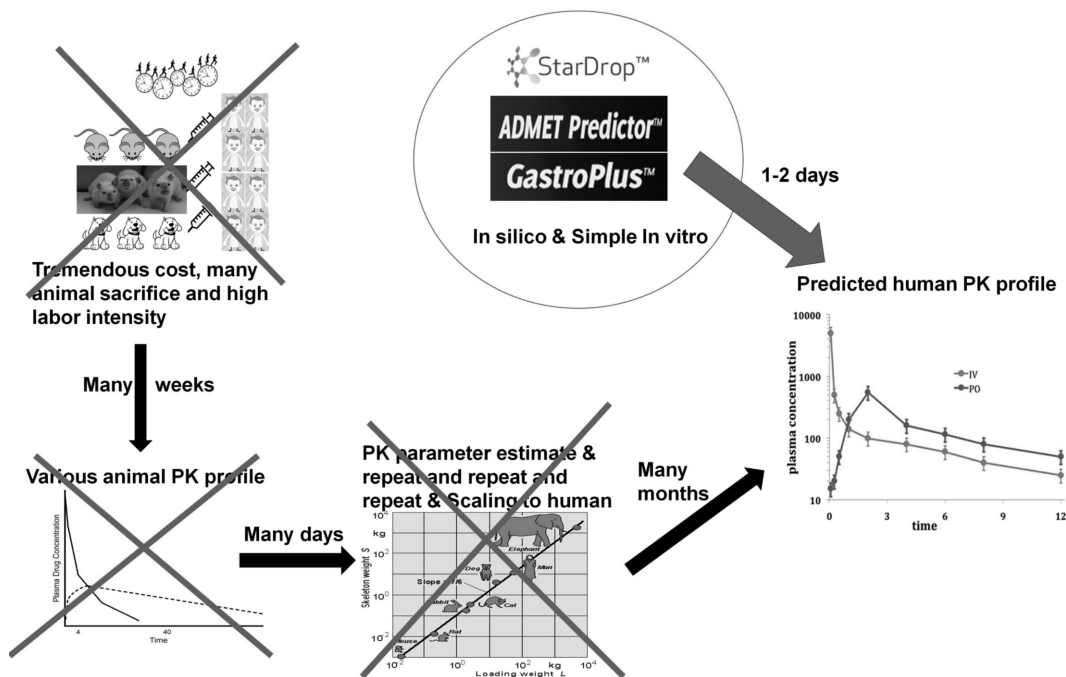


Fig. 8 – Proposed novel strategy of drug discovery.

효율적이면서 빠르게 선정하는 데 큰 역할을 할 것으로 사료된다. 아직은 국내에서 *in silico* 모델링을 이용한 신약 개발 전략이 많이 알려져 있지 않지만 적절한 *virtual library* 합성 전략 및 QSAR를 통한 ADME 검색모델 구현, 약효 예측 모델 구현 및 PBPK 모델링을 통한 동물/인간의 *in vivo* PK 예측 등을 통해서 부족한 자원과 시간 하에 다국적 제약사와 경쟁을 해야 하는 국내 제약사들이 보다 효율적으로 가능성이 있는 신약을 개발하는 데 크게 도움이 될 것이다.

References

- 1) Wenlock, M. C., Austin, R. P., Barton, P., Davis, A. M. and Leeson, P. D. : A comparison of physicochemical property profiles of development and marketed oral drugs. *J. Med. Chem.* **46**, 1250 (2003).
- 2) Gleeson, M. P., Hersey, A., Montanari, D. and Overington, J. : Probing the links between *in vitro* potency, ADMET and physicochemical parameters. *Nat. Rev. Drug Discov.* **10**, 197 (2011).
- 3) Kenny, J. R. : Predictive DMPK: *in silico* ADME predictions in drug discovery. *Mol. Pharm.* **10**, 1151 (2013).
- 4) U.S. Food and drug administration, <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm292923.pdf>.
- 5) Jones, H. and Rowland-Yeo, K. : Basic concepts in physiologically based pharmacokinetic modeling in drug discovery and development. *CPT Pharmacometrics Syst. Pharmacol.* **2**, e63 (2013).
- 6) Grime, K. H., Barton, P. and McGinnity, D. F. : Application of *in silico*, *in vitro* and preclinical pharmacokinetic data for the effective and efficient prediction of human pharmacokinetics. *Mol. Pharm.* **10**, 1191 (2013).
- 7) Ortwine, D. F. and Aliagas, I. : Physicochemical and DMPK *in silico* models: facilitating their use by medicinal chemists. *Mol. Pharm.* **10**, 1153 (2013).
- 8) Hosea, N. A. and Jones, H. M. : Predicting pharmacokinetic profiles using *in silico* derived parameters. *Mol. Pharm.* **10**, 1207 (2013).
- 9) Rueeger, H., Lueoend, R., Rogel, O., Rondeau, J. M., Mobitz, H., Machauer, R., Jacobson, L., Staufienbiel, M., Desrayaud, S. and Neumann, U. : Discovery of cyclic sulfone hydroxyethylamines as potent and selective beta-site APP-cleaving enzyme 1 (BACE1) inhibitors : structure-based design and *in vivo* reduction of amyloid beta-peptides. *J. Med. Chem.* **55**, 3364 (2012).
- 10) 의약화학 편집위원회 : 의약화학, 개정판, 신일북스, 서울 p. 47 (2011).
- 11) Wager, T. T., Hou, X., Verhoest, P. R. and Villalobos, A. : Moving beyond rules: the development of a central nervous system multiparameter optimization (CNS MPO) approach to enable alignment of druglike properties. *ACS Chem. Neurosci.* **1**, 435 (2010).
- 12) Rowland, M., Peck, C. and Tucker, G. : Physiologically-based pharmacokinetics in drug development and regulatory science. *Annu. Rev. Pharmacol. Toxicol.* **51**, 45 (2011).
- 13) Khalil, F. and Läer, S. : Physiologically based pharmacokinetic models in the prediction of oral drug exposure over the entire pediatric age range—sotalol as a model drug. *AAPS J.* **16**, 226 (2014).
- 14) Maharaj, A., Barrett, J. and Edginton, A. : A workflow example of PBPK modeling to support pediatric research and development : case study with lorazepam. *AAPS J.* **15**, 455 (2013).
- 15) Yusof, I., Shah, F., Hashimoto, T., Segall, M. D. and Greene, N. : Finding the rules for successful drug optimisation. *Drug Discov. Today* **19**, 680 (2014).
- 16) U.S. Food and drug administration, http://www.accessdata.fda.gov/drugsatfda_docs/label/2015/205552s0021b1.pdf.