



構造活性相関部会・ニュースレター <1 October, 2010>

SAR News No.19

「目次」

//// Perspective/Retrospective ////

創薬におけるケモインフォマテックス手法の活用 長谷川 清 … 2

//// Cutting Edge ////

創薬プロジェクトにおけるモデリング&シミュレーションの拡張 大川 和史 … 8

Chemoinformatics Toolkits を用いた創薬システム開発における

ラピッドプロトタイピング 吉森 篤史 … 14

//// Activities ////

< 報告 >

構造活性フォーラム 2010 開催報告「QSAR パラダイムの分化と深化」

中川 好秋 … 21

< 会告 >

第 38 回構造活性相関シンポジウム

… 22

///// Perspective/Retrospective /////

創薬におけるケモインフォマテックス手法の活用

中外製薬・鎌倉研究所 長谷川清

1. はじめに

ケモインフォマテックスとは、化学情報全般を扱う広大な学問領域で、データベースの storage, organization, management, retrieval あるいは、データの analysis, dissemination, visualization と種々のテクノロジーを意味する[1]。ケモインフォマテックスは化学構造を基にしたアプローチなので、創薬を目指した製薬企業では、これを利用した薬物設計が盛んである。例えば、部分最小2乗法 (PLS)、サポートベクトルマシーン (SVM) などの統計・機械学習手法は、ケモインフォマテックスおよび関連のケモメトリックスで開発された手法であり、定量的構造活性相関研究やデータマイニング研究に最近多用されている[2]。

創薬の現場においては、特に、ターゲットタンパク質に対するリード化合物の探索およびリード化合物の最適化の効率化が望まれる。本寄稿では、これらの目的のために、いかにケモインフォマテックス手法が有効活用できるかを、最新の研究を紹介しながら説明したい。なお、より広範なケモインフォマテックス手法については、船津らの単行本[1] に詳しくあるので、興味のある読者は一読を薦める。

2. リード化合物の探索

Pharmacophore モデルによる virtual screening

既存の薬物の部分構造を修飾すると、本来のターゲットに対する阻害活性が消失し、かわりに別のターゲットの阻害活性が上昇することが知られている。図1には、ドパミン受容体に阻害活性がある市販の minaprine の側鎖を構造修飾すると、ムスカリン受容体への阻害活性が上昇する実例を示した[3]。

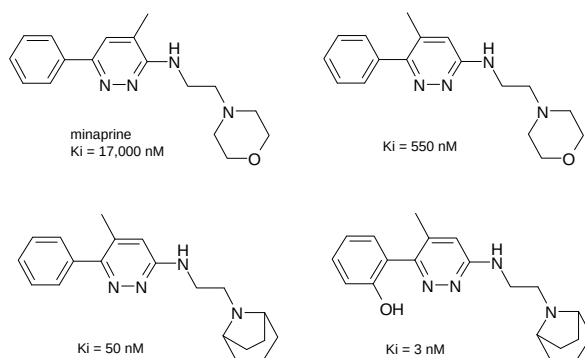


図1 Minaprine の構造修飾によるムスカリン受容体阻害活性の最適化

したがって、ターゲットごとに pharmacophore モデルを用意して、virtual screening を実施すれば、安全性がすでに確保された既存の化学骨格で、別の阻害作用をもつリード化合物を得ることが可能である[4]。個別の pharmacophore モデルからそれぞれの阻害活性を識別できることは、Langer らによりすでに報告があり、このようなアプローチが有用であることが実証されている[5]。図2は、その検証結果である。HIV protease, HIV reverse transcriptase (RT), Influenza neuraminidase (NA), Human rhinovirus (HRV) coat protein, HCV polymerase sites 1,2,3 のターゲットについて、それぞれの pharmacophore モデルがどの程度正しく対応するリガンドを認識できたかを表している。縦軸が pharmacophore モデルの予測、横軸が実測のリガンドである。緑色が正し

く認識された領域を、オレンジ色が誤認識された領域を示す。Influenza NA, HRV coat protein で一部誤認識はあるが、virtual screening という枠組みでは、おおむね正しく認識されていると言える。現在、Langer らは、約 650 個のターゲットについて pharmacophore モデルを用意しており、ターゲットの選択性を考慮した virtual screening が可能である[6]。

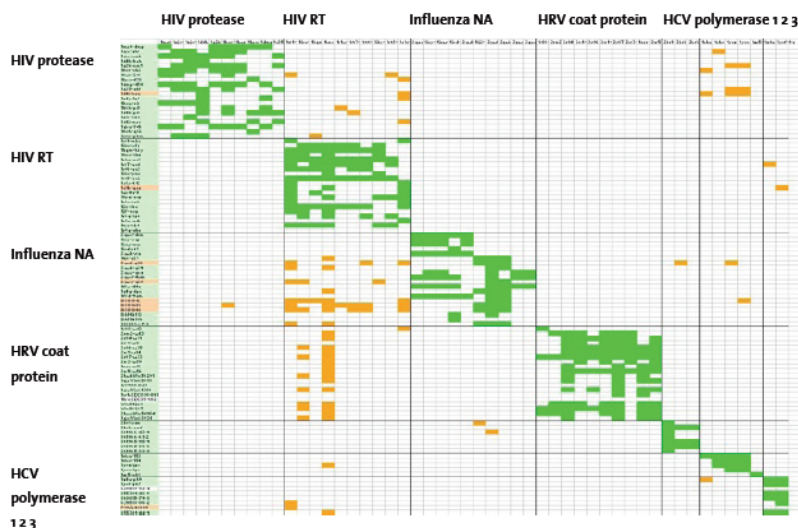


図2 Pharmacophore モデルによる各ターゲットの認識結果 (Langer より許可を得て転載)

De novo design による自動構造発生

既存の阻害剤と化学骨格が異なる化合物を見出すことは、新規特許の点からも分子設計上、重要である。Schneider らは、DOGS (Design of Genuine Structures) という ligand-based de novo design システムを構築して、新規骨格化合物を自動生成することに成功した[7,8] (図3)。まず、あらかじめ vander catalogue にある化合物からユニークなフラグメントを用意しておく。これら fragment を、化学反応を考慮して、任意の組み合わせでつなぐことにより、化学構造が自動発生される。最後に、発生された構造から、ターゲット分子に対する物理化学的特徴の類似性が高い構造を選択する。

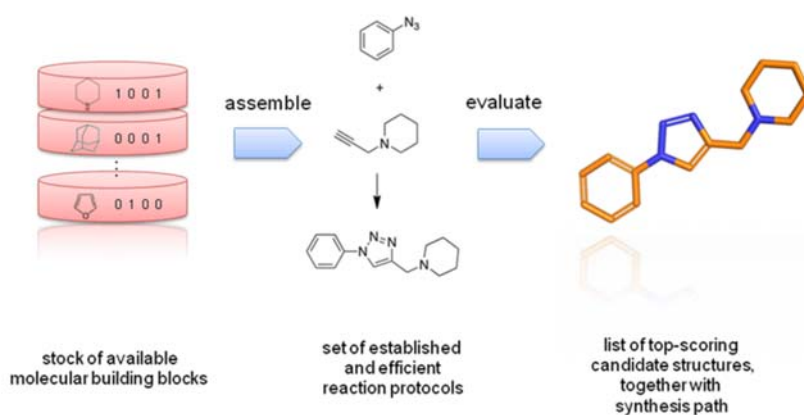


図3 De novo design システム DOGS の概要 (Schneider より許可を得て転載)

TGF- β 1 receptor ligand での実例を図4に示す。既存のリリーの化合物 (LY364947) をテンプレートにして、右側に示すような様々な化合物が出力される。興味深い点は、これまでの de novo design ツールと比較して、出力化合物の合成ルートが明示されている点である。これにより、化

学合成可能なより現実的な化学構造を期待することができる。DOGS は、ligand-based design を志向しているので、タンパク質の構造が未知のプロジェクトでの広い応用が期待される。

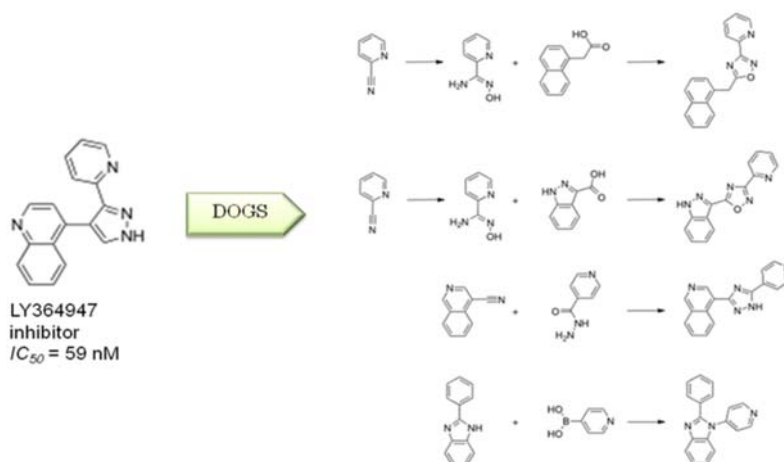


図4 TGF-β1 receptor ligand での実例 (Schneider より許可を得て転載)

3. リード化合物の最適化

ローカルモデルによる ADMET 予測

リード化合物を最適化する場合、阻害活性だけでなく absorption, distribution, metabolism, excretion, toxicity (ADMET) の最適化も必要となる。ADMET 予測プログラムは市販にもあるが、これらは、一般的な化学構造に適した予測精度の低いグローバルモデルである。プロジェクトで合成されているような化学骨格が固定された化合物については、ローカルモデルが適している。長谷川らは、プロジェクト化合物専用の ADMET ローカルモデルとメディシナルケミスト向け専用 web page を社内で構築した[9, 10]。そのイメージ図を図5に示した。

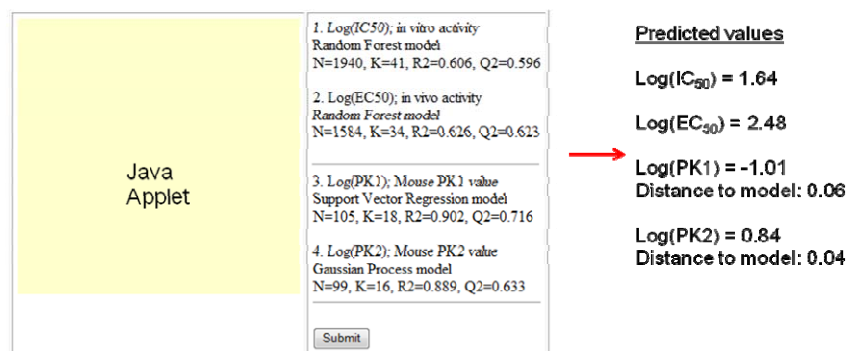
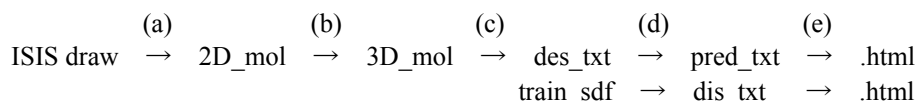


図5 メディシナルケミスト向けプロジェクト専用 web page

Java Applet の部分をダブルクリックすると ISIS/Draw が起動して、任意の化学構造式を入力できる。その後、Submit ボタンを押すと、プロジェクト内で登録されている ADMET 項目の予測値が帰ってくるようになっている。それぞれの予測値には、モデルの適用範囲を示す距離が表示され、この値が大きいと予測値の信頼性が低いことを示す。

予測システムのワークフローを図6に示す。ISIS/Draw で書かれた構造イメージは、Chime Pro により、2次元の mol file に変換される (プロセス(a))。Corina で3次元 mol file に変換された後 (プロセス(b))、MOE で構造記述子が計算される (プロセス(c))。構造記述子が入っている text file から、あらかじめ作成しておいた R モデルで予測値を算出する (プロセス(d))。最後に、予測値を出力 web page の html に cgi で記述する (プロセス(e))。一方、先ほど計算された構造記述子とモデル構築に利用した構造 sdf から、モデル空間までの距離が計算され (プロセス(d))、これもまた出力 web page の html に cgi で記述される (プロセス(e))。



- (a) <applet name="ChimeDecoder" code="ChimeDecoder.class" archive="ChimePro.jar"> </applet>
 (b) corina -i t=sdf input.mol -d wh -d rs -o t=sdf input_opt.mol
 (c) moe/bin/sddesc_list -calc -list deslist.txt -ascii -tab input_opt.mol -o input_opt.txt
 (d) R -f pred.csh → \$pred
 R -f dis.csh → \$dis
 (e) printf ("Log(PK)=%6.2f
", \$pred)
 printf ("Distance=%6.2f
", \$dis)

図6 予測システムのワークフローと具体的なコマンド

専用 web page は、その後いろいろな機能が拡充され、現在、プロジェクト推進に大いに利用されている。追加機能としては、多分子からなる sdf file での予測、分子中のどの部分が重要であるかを示す atom-coloring 表示、予測化合物が chemical space 上どこにあるかを示す図示化機能などがある。

Chemical space の図示化

活性化化合物が chemical space 上にどこにあり、現在合成を考えている化合物がどこにあるかを把握することは、構造最適化ではきわめて重要である。これら全体像がイメージできれば、合成化合物の数を減らすことができかつ最適化のゴールが定量化できる。Bajorath らは、Java を基本とした SARANEA[11] というシステムを開発した。図7は、論文にある plasmin に対する阻害化合物の chemical space を図示化したものである。それぞれのノードが化合物を表しており、Tanimoto 係数>0.65 の化合物ペアは線で結ばれている。赤、オレンジ、黄、黄緑、緑という順番で、阻害活性値の減少を意味している。一方、ノードの大きさは、QSAR cliff の度合いを意味している。化合物ペアで活性値の差が大きい場合、QSAR cliff が大きいと定義される。この値が大きいと chemical space 上断崖絶壁と考えられ、リード化合物の最適化が難しいことを意味している。図7で、continuous と書いてある領域はリード最適化に適しているが、discontinuous と書いてある領域は難しいことを意味する。なお、SARANEA は Java を基本としているので、それぞれのノードにマウスをかざすと、具体的な化学構造と活性データが表示される。

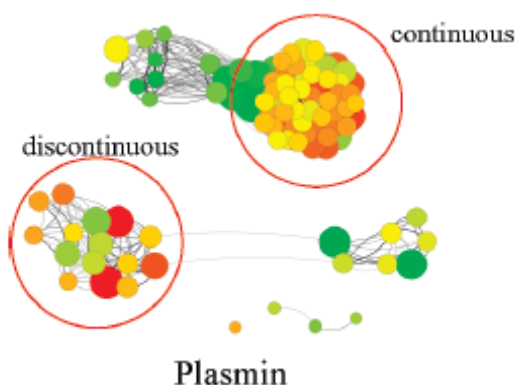


図7 Plasmin に対する chemical space の図示化 (Bajorath より許可を得て転載)

図8は、SARANEA で注目する化合物を中心にして、chemical space 上、どの近接ポイントにどのような化合物があるかを示した図である。中心化合物と比較して、chemical space 上離れれば離れるほど、外側の円上にノードが位置される。円の右側が中心化合物と比較して活性が上昇している化合物を、左側が活性が減少している化合物を示している。このような図から、化合物

をどのように修飾すれば、活性が上昇・減少できるかが一目でわかるようになる。例えば、化合物 1 の右側の ethylamine を化合物 5 の piperidine に変換すれば、活性が上昇することがわかる。近接ポイントだけ表示されているので、骨格部分の構造が固定されデザインのアイデアが容易である。このように、chemical space の図示化は重要なツールであり、先のセクションで記述したローカルモデルと併用すれば、構造最適化がより効率的になる。

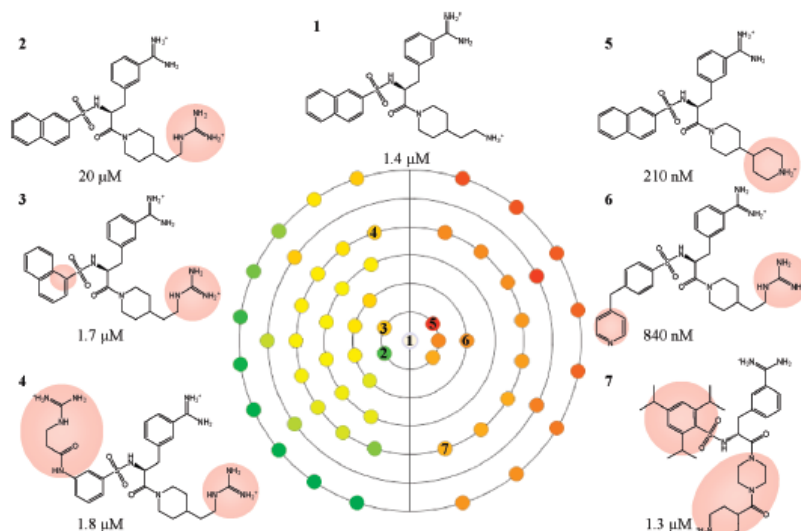


図8 Plasmin に対する chemical space の同心円 (Bajorath より許可を得て転載)

4. おわりに

最近のホットなトピックスとして、ケモゲノミクス研究がある。ケモゲノミクス研究では、多くのターゲットに対する大量のリガンドの活性データを有効に解析し、ラフな傾向を導き出すことが求められている[12]。したがって、ターゲット空間を表現できるバイオインフォマテックスとリガンド空間を表現できるケモインフォマテックスを融合できれば、ターゲットーリガンド空間の効率的な探索が期待できる。さらに、ターゲットーリガンド空間の活性領域内で、構造ジェネレータを利用して構造を発生させれば、オフターゲットに対する新規リガンド化合物を見つけることができるかもしれない[13]。このように、ケモインフォマテックスでは有用なアルゴリズムが順次開発されており、創薬研究に携わる研究者として、今後、この分野の発展を注目していきたい。

参考文献

1. ケモインフォマテックス予測と設計のための化学情報学、船津公人監訳、丸善株式会社 (2005).
2. Chemoinformatics and Advanced Machine Learning Perspectives: Complex Computational Methods and Collaborative Techniques, Lodhi H. and Yamanishi Y. (Eds.), IGI publishing (2011).
3. Aminopyridazines. An Alternative Route to Potent Muscarinic Agonists with No Cholinergic Syndrome: Wermuth C. G., *Il Farmaco*, 48, 253-274 (1993).
4. New Uses for Old Drugs: Chong C. R. and Sullivan D. J., *Nature*, 448, 645-646 (2007).
5. Parallel Screening: A Novel Concept in Pharmacophore Modeling and Virtual Screening: Steindl T. M., Schuster D., Laggner C. and Langer T., *J. Chem. Inf. Model.*, 46, 2146-2157 (2006).
6. Pharmacophores in Drug Research: Langer T., *Mol. Inf.*, 29, 470-475 (2010).

7. Voyages to the (Un) known: Adaptive Design of Bioactive Compounds: Schneider G., Hartenfeller M., Reutlinger M., Tanrikulu Y., Proschak E. and Schneider P., *Trends Biotechnol.*, 27, 18-26 (2009).
8. From Machine Learning to Natural Product Derivatives Selectively Activating Transcription Factor PPAR gamma: Rupp M., Schroeter T., Steri R., Zettl H., Proschak E., Hansen K., Rau O., Schwarz O., Müller-Kuhrt L., Schubert-Zsilavecz M., Müller K.-R and Schneider G., *ChemMedChem*, 5, 191-194 (2010).
9. ADMET ローカルモデルの構築およびケミスト向け Web GUI の開発、長谷川清、深海隆明、大田雅照、白鳥康彦、第 32 回情報化学討論会、山口大学 (2009).
10. Construction of ADMET Local Models and Development of Web GUI for Chemists: Hasegawa K., Fukami T., Ohta M., Shiratori Y., *2nd Chemoinformatics Strasbourg Summer School*, France (2010).
11. SARANEA: A Freely Available Program to Mine Structure–Activity and Structure–Selectivity Relationship Information in Compound Data Sets: Lounkine E., Wawer M., Wassermann A. M. and Bajorath J., *J. Chem. Inf. Model.*, 50, 68-78 (2010).
12. Chemogenomic Approaches to Rational Drug Design: Rognan D., *Br. J. Pharmacol.*, 152, 38-52 (2007).
13. Exhaustive Structure Generation for Inverse-QSPR/QSAR: Miyao T., Arakawa M. and Funatsu K., *Mol. Inf.*, 29, 111-125 (2010).

//// Cutting Edge ////

創薬プロジェクトにおけるモデリング&シミュレーションの拡張

持田製薬株式会社 創薬研究所 化学研究室
大川和史

1. はじめに

化合物の置換基変換を行ったり、反応試薬データベースを使った仮想的な構造を創出するようなケモインフォマティクスツールはポピュラーであるが、まずは基本に立ち戻って化学構造の変換をグラフ操作の観点から捉えなおす。次に、これらの操作を拡張することにより、ケミストが合成案を創出するプロセスを、反応のルールセットの集合とその操作を用い表現することを考える。さらにそれらを束ね拡張した場合にプロジェクトが表現できるであろうことを示す。次の章では、構造最適化プロジェクトというものを、ストリームとして捉えてみる。その場合に、化合物管理システムは単に化合物のデータを管理するだけではなく、よりメタ情報（情報に関する情報）を記録していく必要があるが、ソフトウェア開発のプロジェクトマネジメントシステムではすでにそのようなメタ情報を管理するようなシステムが存在するため、それらと比較しながら、メタ化合物管理システムというものを考えてみたい。

2. 化学構造プログラミング

化学構造はノードとエッジで表現できるため、いくつかのケモインフォマティクス用のライブラリではグラフの操作として化学構造を扱うことができる。そのようなライブラリとして OEChem Toolkit[1]、PerlMol[2]、OpenBabel[3]が存在する。このようなライブラリはグラフ操作を通して化学構造を修飾できるが、化学反応を記述するためのフォーマット（例えば SMIRKS[4]、RXN[5]）フォーマットをサポートしていれば煩雑なグラフ操作をしなくても構造の変換が可能である。OEChem Toolkit や Pipeline Pilot[6]では RXN フォーマットがサポートされているのでグラフ操作の処理を自分で書く必要はないが、本稿ではグラフ操作のイメージを持ってもらうため PerlMol によるコードを二つ載せる。

例 1) Diels-Alder 反応

例えば PerlMol では SMARTS[7]というパターンマッチのための表現形式を使うことで化学反応を定義できる。例として図 1 のような Diels-Alder 反応を記述してみる。

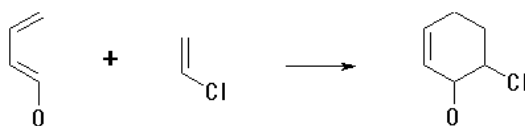


図 1 Diels-Alder Reaction

```
use Chemistry::File::SMILES;
use Chemistry::File::SMARTS;
use Chemistry::Ring 'aromatize_mol';

my $react1 = Chemistry::Mol->parse('ClC=C dienophile',format=>'smiles');
my $react2 = Chemistry::Mol->parse('OC=C-C=C diene',format=>'smiles');

#diene と dienophile のパターンを SMARTS で定義
my $dienophile_pat = Chemistry::Pattern->parse('C=C',format=>'smarts');
my $diene_pat = Chemistry::Pattern->parse('C=C-C=C',format=>'smarts');
```

```

aromatize_mol($react1);aromatize_mol($react2);

my $name = $react1->name."+ ".$react2->name;
my $prod = Chemistry::Mol->new(name=>$name);
$prod->combine($react1,$react2);

#diene と dienophile のパターンを探し、みつかったら各オブジェクトにマップ
$dienophile_pat->match($prod);
my @atom_map1=$dienophile_pat->atom_map;
my @bond_map1=$dienophile_pat->bond_map;

$diene_pat->match($prod);
my @atom_map2=$diene_pat->atom_map;
my @bond_map2=$diene_pat->bond_map;

#反応用の原子が揃っているかチェック
if($atom_map1[0] && $atom_map1[1] && $atom_map2[0] && $atom_map2[3]){

#結合の作成
$prod->new_bond(atoms=>[$atom_map1[0],$atom_map2[0]],order=>'1');
$prod->new_bond(atoms=>[$atom_map1[1],$atom_map2[3]],order=>'1');

#diene と dienophile の結合次数を変える
$bond_map1[0]->order(1);$bond_map2[0]->order(1);
$bond_map2[1]->order(2);$bond_map2[2]->order(1);

my $smi=$prod->print(format=>'smiles',unique=>1,name=>1);
print "$smi\n";
}

```

例 2) ベンゼンを生物学的等価なチオフエンに変換する

コンピュータ上の操作であるため、グラフ操作であればどのような構造の修飾でも可能である。例えば生物学的等価な変換として知られているベンゼンをチオフエンに変換する(図 2)コードは以下のようになる。

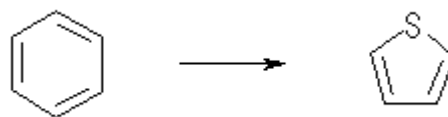


図 2 ベンゼンをチオフエンに

```

use Chemistry::File::SMILES;
use Chemistry::File::SMARTS;
use Chemistry::Ring 'aromatize_mol';

my $comp = Chemistry::Mol->parse('c1ccccc1',format=>'smiles');
aromatize_mol($comp);
my $pattern = Chemistry::Pattern->parse(
    'c:[ch]=[ch]:c:c:c',format=>'smarts');

$pattern->match($comp);
my @nfg = $pattern->atom_map;

$nfg[1]->delete; $nfg[2]->delete;

my $s = $comp->new_atom(symbol=>"S");
my $sc1 = $comp->new_bond(atoms=>[$s,$nfg[0]],order=>'1');
my $sc2 = $comp->new_bond(atoms=>[$s,$nfg[3]],order=>'1');
$s->implicit_hydrogens(0);

print $comp->print(format=>'smiles',unique=>1,name=>0),"n";

```

Bioisostere generator

ここまでで化学構造の変換をプログラミングできることを示したが、これをさらに広げていこう。例えば文献などで経験上知られている生物学的等価体のルールをすべて収集し、任意の化合物に対してすべての可能な生物学的等価体を生成するようなツールを用意すれば、ケミストの経

験や知識のばらつきによらずに、ボトムラインとしての一定の質が担保できる。弊社ではこのようなツールを web サービスとして公開することで、ケミストが容易に興味のある構造からすべての可能な生物学的等価体を SDF ファイルとして受け取れるようにしている。他の事例としては、aromatic ring に再帰的にヘテロ原子を入れていき、すべての組み合わせを自動的に出力する、MORPH[8]のようなツールも有用性が高いと思われる。

ケミストを模倣する

より複雑な拡張として、ケミストのように合成案を創出するような仕組みを考える。例えば、環を切ったり巻いたりといった大胆な骨格変換、アミド結合をリバースにしたりといった変換から、単にアルキル基、ハロゲン基の導入といった細かい置換や、上で挙げた生物学的等価な変換やヘテロ環の組み合わせをランダムに生成していくようなプログラム[9]は可能である。

単にランダムに構造を生み出したいだけであれば、反応ルールのセットを保持しておいて、反応ルートをランダムに選択しながら新たな構造を生成、新しく生まれた構造を再度反応ルールにかけるような再帰的なループを回せばよい(図 3)が、実用性を考えた場合、反応ルールセットは設定ファイルで調整できるようにし、新規な化合物のプロパティを評価しながら、イテレーションごとに設定ファイルを調整するような仕組みにしたほうがよいであろう。そうすれば「極性が高いようであれば次のイテレーションではそれを下げる」「分子量が大きくなった場合、次のイテレーションでは構造を削る」ような反応ルールが選択されやすくなるように設定ファイルを書き換えることができる。

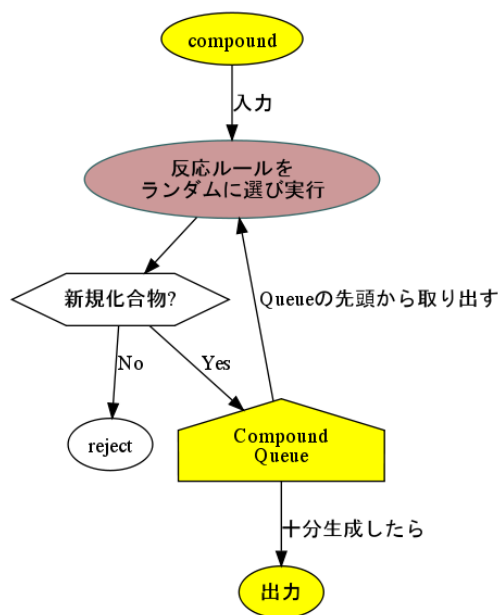


図 3 ランダム合成サイクル

さらに、このシミュレーションの大きな利点の一つは、世代を管理することができ、どの構造から生まれた構造かが明確になっているために、時系列を考慮したさらに大きいシミュレーションの枠組みとして使える可能性があるということである。リードオプティマイゼーションは単に化合物を積み上げていくだけでなく、プロジェクトの歴史を刻みながら、プロジェクトそのものを科学的に理解していく作業であると筆者は考えている。次の章ではこのような世代を扱えるシミュレーションツールを使ったプロジェクトシミュレーションの可能性を考えてみたい。

3. プロジェクト

プロジェクトの構造

リードオプティマイゼーションのプロセスでは、構造最適化されるべき出発化合物が一つまたは複数存在するが、これらの化合物が修飾された後、さらに新しく合成された化合物をもとに新たな化合物が合成されるため、最終的にはプロジェクトの出発化合物をルートとしたツリー構造を形成することになる。ここで、親の化合物に対して子を幾つか合成してみて、そのうちの特性の良いものの上位数化合物から新たな子を生成する(ただし最初のほうは特性が振れやすく最終的には安定する)というようなモデルをつくってシミュレーション[10]してみると以下のような図になる。グラフの描画は Cytoscape[11]を用いたが、このような情報もまたグラフとして表現できる。図 4 において、ノードに化合物、エッジには親子関係をあらわした。世代を経るにつれエッジの色が黄色から緑に変化するようにし、ノードの色は活性のようなものを表現しており、黒から赤になるに従い特性が良くなるようにした。

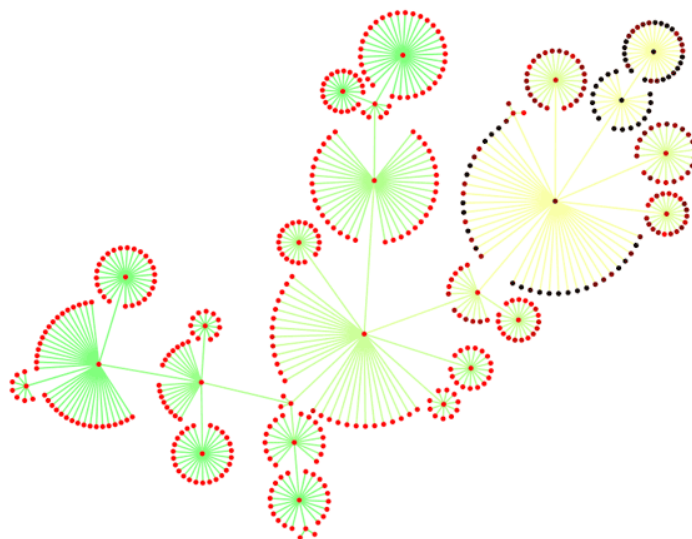


図4 プロジェクトにおける化合物のシミュレーション

実際のプロジェクトにおいては、親子関係がもっと複雑であったり、解決すべき課題が複数あるためにこのような単純な構造は取らないだろうが、プロジェクト全体の「見える化」としては有用な表現方法だと考えられる。

プロジェクトシミュレーション

John Delaney[12]は random-walk の一種である self-avoiding walk を用いて化合物最適化のシミュレーションを行ない、プロジェクトの成功確率と投入する人員の関係性や、ラボの生産性の考察を行っており、その結果は興味深い。しかし、本論文においては self-avoiding walk を用いたシミュレーションであるために、化合物の構造変換を陽には考慮していない。そこで、先にあげた、ケミストを模倣するようなシミュレーションの枠組みを用意することで化合物の構造を明示的に取り入れ、プロジェクトシミュレーションをより具体的なレベルで行いたいというのが筆者の考えである。もし、このような枠組みを用意できれば、実際のプロジェクトと比較しながらよりよいプロジェクトの意思決定に貢献するようなシステムを構築できるであろうと考えている。

もちろん、このようなシミュレーションの評価のためには実際のプロジェクトの歴史が記録されたデータが必要である。そこで、最後の章では、そのようなプロジェクトマネジメントシステムの可能性に関して考えてみる。

4. メタ化合物管理システム

プロジェクトマネジメントサイクルにおいてはソフトウェア開発と構造最適化プロジェクトに多くの共通点が見られるため、ソフトウェアプロジェクトマネジメントシステムのアナロジーから考えてみる。実際にソフトウェア開発においては既に幾つものアプローチが試みられていて実装も多い。

ソフトウェアのバージョン管理システムとしては Subversion[13]、Git[14]、Mercurial[15]といったツールが有名である。これらの基本的な機能は、ファイルの作成日時、変更日時、変更点などの履歴を保管することである。修正したソフトウェアをリポジトリに追加する行為をコミットまたはチェックインといいコミットを積み重ねることでそのプロジェクトの歴史が形作られ、この歴史のことをコミットツリー(図5)などと呼ぶ。

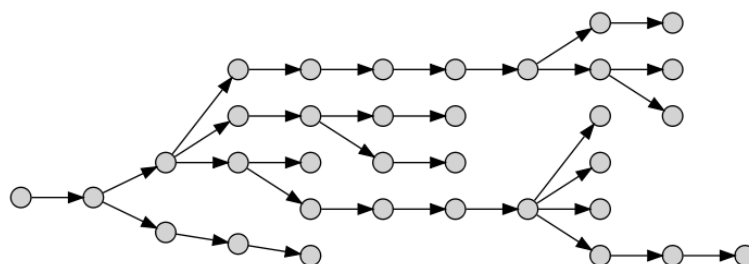


図5 コミットツリー

バージョン管理システムでは複数の人間が扱うことを想定した機能になっており、どのソースコードを改編したかを記憶している。さらに、ソースコードのどの部分が改編されているかといった親との差分が簡単に見えるようになっている。

いまここで、そのような化合物管理データベースを考えてみよう。つまりどの化合物から着想を得て改変したかを記録してあるデータベースマネジメントシステムが存在し、いつ、誰がどのような化合物をどの化合物に基づいて合成したかを記録できているとすれば、コミットツリーの矢印は差分を示しており

- 構造のどの部分が変化したのか
- cLogP、PSA や実際に測定した物性値がどのように変化したのか
- 実際に測定したアッセイ系での変化量

といった差分情報が記録されることになるだろう。バージョン管理システムにおいては、差分が簡単に把握できるため、自分が合成した化合物が何を变化させたのか？そしてその結果プロジェクトにどう影響をあたえたのか？がより明確になる。つまり親子の構造の差分はすなわちケミストの意図やプロジェクトの意思決定結果を反映しており、それらを「見える化」することには大きな意味があるわけである。モデリングの立場からも構造の差分は反応ルールに還元するための重要な情報であるし、このようにデータベース化されていれば、究極的には知識としての反応ルール抽出の自動化が可能になり、シミュレーターにもフィードバックできるようになるため、有用である。

さらにマネジメントシステムには並行して走る複数の内部プロジェクトを管理する仕組みが備わっており、違うスキマフォルドを並行して複数展開していく場合にも適切に管理できる。進捗具合や有害事象またはバックアップ化合物の構造最適化が行なわれたり、初期の構造に戻りなおし、新たに合成展開をやり直すといった状況でも、きちんと歴史に残せる。

一方、このようなアナロジーでもう一つ見えてくるものは、創薬系のプロジェクトマネジメントにおいてバグトラッキングシステムやイシュー管理システムのようなものが見られないことである。デザインなどのクリエイティブ要素を含む開発現場でも最近をよく使われているという状況において、ドラッグデザインといった、かなりクリエイティブ性の高い領域であり使われていない状況に筆者は違和感を感じる。ソフトウェアプロジェクトマネジメントシステムには Trac[16]や Redmine[17]といったバグトラッキングシステムが有名であるが、このように発生した要求変更やバグと歴史の任意の時点を紐づけるような管理の手法が創薬プロジェクトにおいて、今後必要とされるのではないだろうか？

5. おわりに

モデリング&シミュレーションの大きな目的のひとつはモデリングの作業を通して現実への理解を深めることである。つまり、作ったモデルと実際のデータのギャップデータは何か、隠れた重要な特徴は何か、それを含めた場合に観測されたデータをより特徴づけられるかといったトライアンドエラーを繰り返すことで、よりよい意思決定を行なうのである。

プロジェクトのある時間軸(ブランチの群)を用いて SAR を行った結果に対し、実際にこれから合成する化合物は interpolation するのか extrapolation するのかきちんと考えることは実際のプロジェクトにおいては非常に重要であり、これらはプロジェクトがどういう歴史をたどってきてこれからどう向かおうとしているのかというメタ情報に影響されるものである。

これらのことを考えると、創薬プロジェクトにおいて、構造活性相関解析が行われる化合物のリストのメタ情報の解析といったプロジェクトそのものを評価していくことは意味があると思われるし、そういった情報を記録するようなデータベースまたはインフラストラクチャの整備は重要であると考えている。

参考文献

- [1] <http://www.eyesopen.com/>
- [2] <http://www.perlmol.org/>
- [3] Guha R, Howard MT, Hutchison GR, Murray-Rust P, Rzepa H, Steinbeck C, Wegner J, Willighagen EL. The Blue Obelisk-interoperability in chemical informatics. J Chem Inf Model. 2006 May-Jun;46(3):991-8.
- [4] http://www.daylight.com/dayhtml_tutorials/languages/smirks/index.html
- [5] <http://www.symyx.com/downloads/public/ctfile/ctfile.jsp>
- [6] <http://accelrys.co.jp/products/pipeline-pilot/>
- [7] <http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html>
- [8] Beno BR, Langley DR. MORPH: a new tool for ligand design. J Chem Inf Model. 2010 Jun 28;50(6):1159-64.
- [9] <http://github.com/kzfm/Chemicho>
- [10] <http://gist.github.com/574635>
- [11] <http://cytoscape.org/index.php>
- [12] Delaney J. Modelling iterative compound optimisation using a self-avoiding walk. Drug Discov Today. 2009 Feb;14(3-4):198-207. Epub 2008 Dec 8.
- [13] <http://subversion.tigris.org/>
- [14] <http://git-scm.com/>
- [15] <http://mercurial.selenic.com/>
- [16] <http://trac.edgewall.org/>
- [17] <http://www.redmine.org/>

///// Cutting Edge /////

Chemoinformatics Toolkits を用いた創薬システム開発における ラピッドプロトタイピング

株式会社理論創薬研究所 吉森篤史

1. はじめに

近年、*in silico* における創薬技術は、*in vitro/in vivo* における技術と同様に、創薬プロセスにおいて必要不可欠な技術となっている。創薬プロジェクトにおいては、内在性リガンド、標的タンパク質、特許情報など、開始時(参加時)に得られている情報がプロジェクトごとに大きく異なるため、それに伴い、Computational Chemist に求められる貢献範囲も様々なものとなる。従って、創薬プロジェクトを成功に導くために *in silico* で何ができるのかを、与えられた情報を最大限に生かし、かつ有益なアウトプットを出せる戦略として、迅速かつ具体的に提示することが、創薬プロジェクトにおける Computational Chemist の重要な役割の1つであると言える。特に迅速性は、創薬プロジェクトにおいて *in silico* に期待される普遍的な要求であると考えられる。

ソフトウェア開発におけるプロトタイピングとは、はじめから完成版といえるソフトウェアの開発に着手するのではなく、まずは、プロトタイプ(試作品)を作成し、完成版のイメージをユーザーに抱かせることにより、初期段階から適切なフィードバックを得るための手法である。創薬プロジェクトにおいては、過去の経験も含め、様々なアイデアをもった研究者が参加しており、プロトタイプ、及びそれから得られるアウトプットの提示は、ソフトウェアの修正、及び改善点をより的確に把握するための重要な工程であると考えられる。

それでは、創薬システムの開発において、迅速なプロトタイピング(ラピッドプロトタイピング)を実現するためには、何が必要であろうか？現在、多くの Chemoinformatics Toolkits が、オープンソースソフトウェアとして公開されている。これら Toolkits には、創薬システムの開発において必要とされるアルゴリズムの多くが実装されており、これらを効率的に利用することで、プロトタイプの開発期間を大幅に短縮することができる。しかしながら、その Application Program Interface (API)は膨大であり、またドキュメントの整備もあまり進んでいないことから、ソースコードを通して、その使用法を理解しなければならないケースも多い。従って、ラピッドプロトタイピングの実現には、適切な Toolkits の選択だけではなく、API の使用法や実践的な Tips の整備が重要な課題であると言える。

本稿では、ラピッドプロトタイピングの適用例として、Chemoinformatics Toolkits の1つである RDKit (<http://www.rdkit.org/>)を用いた標的タンパク質指向型フラグメントライブラリの構築について紹介する。ラピッドプロトタイピングの面白さと有効性をご理解頂ければ幸いである。

2. Chemoinformatics Toolkits

現在、オープンソースソフトウェアとして公開されている代表的な Chemoinformatics Toolkits を表1に示す。

表 1 Chemoinformatics Toolkits

名称	URL	言語
OpenBabel	http://openbabel.org/	C++ / Python 他
JOELib/JOELib2	http://www.ra.cs.uni-tuebingen.de/software/joelib/	Java
CDK	http://sourceforge.net/apps/mediawiki/cdk/	Java
RDKit	http://www.rdkit.org/	C++ / Python
PerIMol	http://www.perlmol.org/	Perl
Cinfony	http://code.google.com/p/cinfony/	Python
OASA	http://bkchem.zirael.org/oasa_en.html	Python

多くの Toolkits は、C++、Java 等で実装されているが、Simplified Wrapper and Interface Generator (SWIG)、Jython、Ruby Java Bridge (RJB)などを通じて、様々なスクリプト言語からの利用が可能となっている。この様な仕組みにより、まずは、スクリプト言語でプログラム全体を記述し、性能評価を行った上で、処理速度のパフォーマンスの改善に必要な部分だけ、オリジナルの実装言語で書くなどの作業を行うことができる。ラピッドプロトタイピングにおけるスクリプト言語の利用は、迅速なシステム開発を実現するための必要不可欠な要素であると言える。

一方、各々の Toolkits は、互いに共通した機能を持っている反面、Fingerprint や構造生成等、異なった機能も併せ持っている。従って、複数の Toolkits を補完的に利用することにより、目的とするシステムの開発をより効率的に行うことができる。しかしながら、Toolkits 間におけるデータの互換性や相互変換の機能はほとんどなく、プログラミングを行う際の大きな障害となっている。Cinfony では、この様な問題の解決策の1つとして、OpenBabel、CDK、及び RDKit を共通のインターフェースを通して利用できる仕組みを提案している[1]。Chemoinformatics Toolkits の利用において、共通インターフェースの普及は、迅速、効率的な創薬システムの開発に大いに貢献するであろう。

3. 適用例：標的タンパク質指向型フラグメントライブラリの構築

近年、Fragment-Based Drug Discovery (FBDD)には、実践的な創薬手法の1つとして、多くの関心が寄せられている。その背景として、活性が弱いながらも適切な結合を形成できる小分子を起点とすることにより、High Throughput Screening (HTS)で得られるリード化合物と比較して、より効率的な合成展開が期待されることが挙げられる。また、FBDDを支える基盤技術として、表面プラズモン共鳴、X線結晶構造解析、核磁気共鳴に代表される物理化学的手法が、目覚ましい技術進歩を成し遂げたことも挙げられる。一方、FBDDを行うに当たり、フラグメントライブラリのデザインは、FBDDの成功のカギを握る重要なステップであると言える。一般的に、フラグメントとして適した分子は、Rule of Three [2]に代表されるフラグメント指標により選別され、ライブラリの構築が行われる。

本稿では、ラピッドプロトタイピングの適用例として、標的タンパク質指向型フラグメントライブラリの構築について紹介する。ここでは、Cell Division Protein Kinase 2 (CDK2)に対するFBDDを仮想的な創薬プロジェクトとして設定した。そして、我々に与えられた課題を、CDK2を対象としたフラグメントライブラリの構築と設定して説明を行う。

3.1 創薬システムの概要

ラピッドプロトタイピングの対象となる創薬システムの概要を図1に示す。

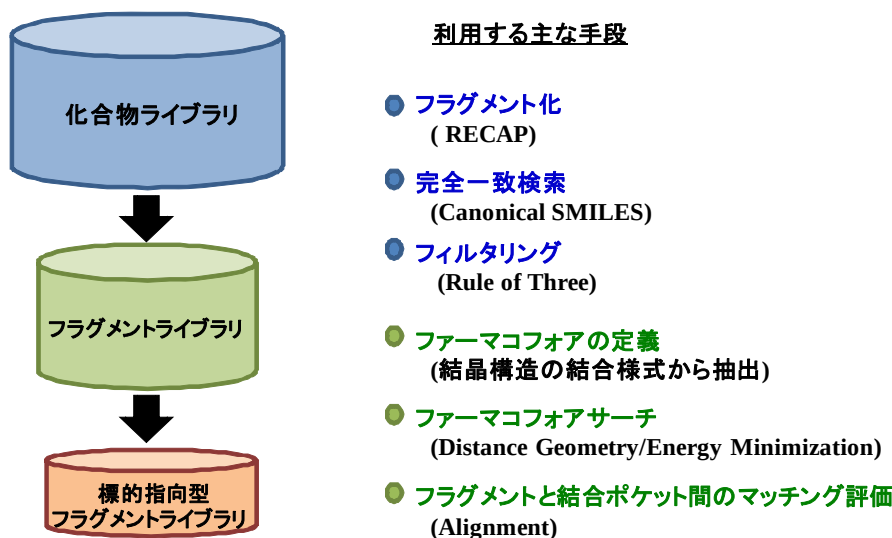


図1 創薬システムの概要

本創薬システムは、はじめに、1) 化合物ライブラリの選別、2) 化合物ライブラリのフラグメント化、3) 重複フラグメントの削除、4) フィルタリングの実施により、フラグメントライブラリを構築する。次に、5) CDK2 と阻害剤の共結晶構造からファーマコフォアの定義、6) ファーマコフォアサーチ、7) フラグメントと結合ポケット間のマッチング評価により、標的タンパク質指向型フラグメントライブラリを構築する。

ラピッドプロトタイピングを行うに当たり、Chemoinformatics Toolkits として RDKit を選択し、言語は Python を使うことにした。ここでは、図 1 の () 内に記載した方法により、本創薬システムのラピッドプロトタイピングを実施した。

3.2 フラグメントライブラリの構築

RDKit では、Retrosynthetic Combinatorial Analysis Procedure (RECAP) [3] により分子のフラグメント化を行うことができる。RECAP は、11 個の切断ルールに基づいて分子をフラグメントに分解する。例えば、切断ルールの 1 つであるアミド結合の切断は、以下のように定義されている (Recap.py, Line 69)。

```
[C;!$(C([#7])([#7]):1)(=[O:2])!@([#7;+0;!D1:3]>>[*][C:1]=[O:2].[*][#7:3]
```

また、Python の対話モードにおいて (RDKit はインストール済みであること)、

```
>>> from rdkit.Chem import Recap
>>> print Recap.reactionDefs
```

により、全ての切断ルールを確認することができる。Recap.py 最新版では、11 個の切断ルールの他に、aromatic nitrogen - aromatic carbon 間の切断ルールが新たに追加されていることが分かる。

RECAP を図 2 の化合物に一般的に適用した場合、1~4 の 4 個のフラグメントに分解される (図 2)。一方、RDKit に実装されている RECAP では、階層的にフラグメント化を行うことができ、この処理により、1~10 の 10 個のフラグメントを得ることができる (図 2)。

```
>>> from rdkit import Chem
>>> from rdkit.Chem import Recap  ↓ 分子を SMILES で入力
>>> mol = Chem.MolFromSmiles("CCCC(=O)Nc1n[nH]c2c1cc(-c1ccccc1)c(-c1ccc(O)cc1)c2")
>>> root = Recap.RecapDecompose(mol) ← RECAP によるフラグメント化
>>> print root.GetLeaves().keys()    ← RECAP の一般的な適用
>>> print root.GetAllChildren().keys() ← RECAP の階層的な適用
```

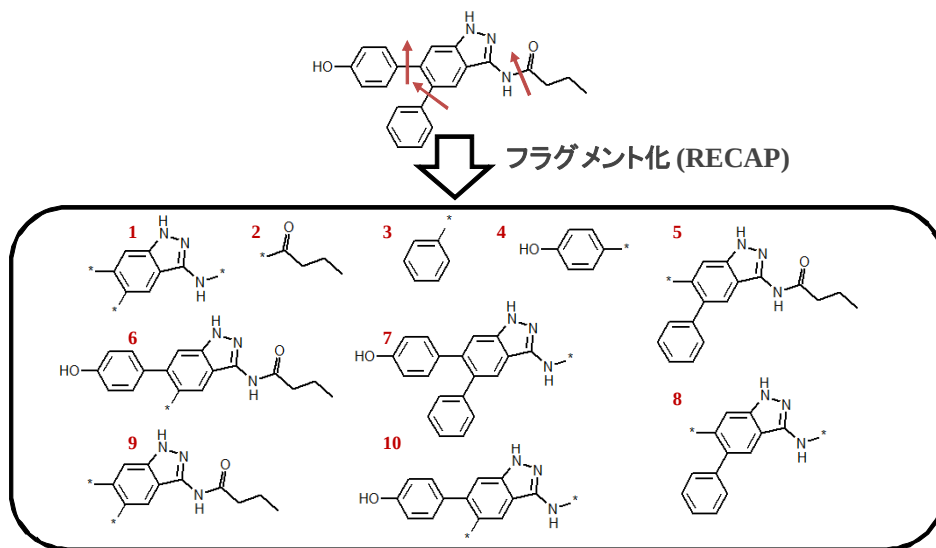


図 2 RECAP を用いたフラグメント化

RECAP を一般的に適用した場合、分子量 100 以下のフラグメントが大部分を占め、また得られるフラグメント数も少ないが、RECAP を階層的に適用した場合、分子量の幅も広く、また多様なフラグメントを得られることが分かった。本稿では、RECAP を階層的に適用して、分子のフラグメント化を行なった。また、切断部位の原子(図 2 の*)は、炭素原子でキャッピングを行い、そして、Canonical SMILES に変換後、文字列比較により、重複の除去を行った。

次に、Rule of Three を用いて、RECAP でフラグメント化された分子のフィルタリングを行った。Rule of Three とは、一般的に、分子量 ≤ 300 、 $\text{cLogP} \leq 3$ 、水素結合供与体数 ≤ 3 、水素結合受容体数 ≤ 3 、回転可能結合数 ≤ 3 、極性表面積 $\leq 60 \text{ \AA}^2$ として定義されたフラグメントの指標である。本稿では、RECAP によるフラグメント化の際、分子量 100 以下のフラグメントが多数含まれること、及び後で定義するファーマコフォアのサイズも考慮して、分子量の下限を設けた $120 \leq \text{分子量} \leq 300$ を採用した。

Rule of Three で用いる記述子は、RDKit の AvailDescriptors モジュールを用いて計算した。

```
>>> from rdkit import Chem
>>> from rdkit.Chem import AvailDescriptors
>>> mol = Chem.MolFromSmiles("CNc1n[nH]c2cc(C)c(C)cc21") ←分子を SMILES で入力
>>> mw = AvailDescriptors.descDict['MolWt'](mol)      #記述子の計算
>>> logP = AvailDescriptors.descDict['MolLogP'](mol)
>>> hd = AvailDescriptors.descDict['NumHDonors'](mol)
>>> ha = AvailDescriptors.descDict['NumHAcceptors'](mol)
>>> rb = AvailDescriptors.descDict['NumRotatableBonds'](mol)
>>> tpsa = AvailDescriptors.descDict['TPSA'](mol)
>>> print mw,logP,hd,ha,rb,tpsa      ↓ Rule of Three によるフィルタリング
>>> if mw >= 120 and mw <= 300 and logP <=3.0 and hd <=3 and ha <=3 and rb <=3 and tpsa <=60:
>>> ...     print "pass"
```

ちなみに、AvailDescriptors モジュールで利用可能な記述子の種類は、176 種類あり、

```
>>> print AvailDescriptors.descList
```

により、確認することができる。

3.3 ファーマコフォアの定義

RDKit では、対象とするリガンド分子内に局在する特性基とその距離に基づき、ファーマコフォアの定義を行うことができる。特性基として、“Donor”、“Acceptor”、“NegIonizable”、“PosIonizable”、“Aromatic”等が BaseFeatures.fdef ファイル内に定義されており、特性基の変更、及び追加も行うことができる。まずは、リガンド分子内の特性基の探索を行うことにした。

```
>>> import os.path
>>> from rdkit import Chem
>>> from rdkit import RDConfig
>>> from rdkit.Chem import ChemicalFeatures      ↓分子を SMILES で入力
>>> mol = Chem.MolFromSmiles("CCCC(=O)Nc1n[nH]c2c1cc(-c1ccccc1)c(-c1ccc(O)cc1)c2")
>>> dataDir = os.path.join(RDConfig.RDCodeDir,'Chem\Pharm3D\test_data')
>>> featFact = ChemicalFeatures.BuildFeatureFactory(os.path.join(dataDir,'BaseFeatures.fdef'))
>>> feats = ChemicalFeatures.MCFF_GetFeaturesForMol(featFact,mol)
>>> for i in range(len(feats)):                    ↑ 分子内にある特性基の検索
>>> ...     family = feats[i].GetFamily()      ←検索された特性基の種類の取得
>>> ...     print family
```

本手順により、分子内に 13 個の特性基が検出された。本分子は、CDK2 との共結晶が Protein Data Bank(<http://www.pdb.org/>)に登録されていることから(PDB ID: 3LFN)[4]、目視により、この 13 個の特性基の中から CDK2 との結合に重要な特性基の抽出を行った(図 3)。

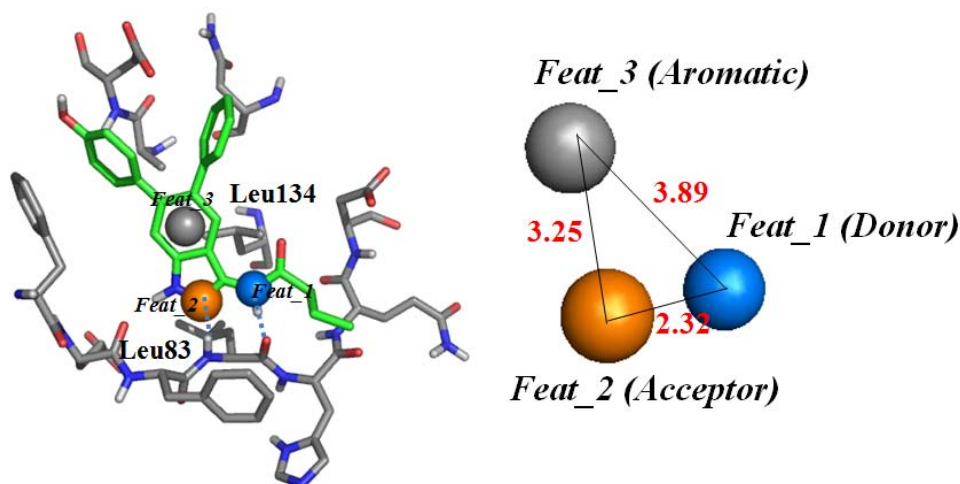


図3 ファーマコフォアの定義

CDK2 と阻害剤の結合様式から、CDK2 のヒンジ領域に位置する Leu83 の=O との水素結合を示す *Feat_1(Donor)*、Leu83 の NH との水素結合を示す *Feat_2(Acceptor)*、そして Leu134 を中心とした疎水領域との結合を示す *Feat_3(Aromatic)* の 3 つの特性基をファーマコフォア形成する特性基として定義した(図 3 左)。そして、各特性基間の距離は、結晶構造から算出した値(単位はÅ)を利用した(図 3 右)。

RDKit におけるファーマコフォアの設定は、以下の手順で実施することができる。

```
>>> from rdkit.Chem import ChemicalFeatures
>>> from rdkit.Chem.Pharm3D import Pharmacophore
>>> from rdkit import Geometry
>>> feat_1=ChemicalFeatures.FreeChemicalFeature('Donor',Geometry.Point3D(43.1,-11.0,23.8))
>>> feat_2=ChemicalFeatures.FreeChemicalFeature('Acceptor',Geometry.Point3D(41.8,-11.2,21.9))
>>> feat_3=ChemicalFeatures.FreeChemicalFeature('Aromatic',Geometry.Point3D(39.8,-9.1,23.2))
>>> feats = [feat_1,feat_2,feat_3]
>>> pcophore = Pharmacophore.Pharmacophore(feats)
>>> d_upper = 1.5
>>> d_lower = 0.5
>>> pcophore.setLowerBound(0,1,2.32-d_lower)
>>> pcophore.setUpperBound(0,1,2.32+d_upper)
>>> pcophore.setLowerBound(1,2,3.25-d_lower)
>>> pcophore.setUpperBound(1,2,3.25+d_upper)
>>> pcophore.setLowerBound(0,2,3.89-d_lower)
>>> pcophore.setUpperBound(0,2,3.89+d_upper)
>>> print pcophore
```

ここでは、便宜上、1 つの共結晶構造からファーマコフォアを定義したが、ファーマコフォアの抽出法には様々な手法があるため、上に示すような設定方法を理解していれば、応用展開も容易と思われる。

3.4 ファーマコフォアサーチ

一般的なファーマコフォアサーチにおいては、分子の 3D 構造生成後、ファーマコフォアのサーチを行うが、RDKit においては、事前に分子の 3D 構造を生成させるのではなく、ファーマコフォアを満たす 3D 構造が生成できるか否かを判定基準として、ファーマコフォアのサーチを行うことができる。

RDKit におけるファーマコフォアサーチの流れを以下に示す。なお、紙面の都合上、重要な部分のみを抜粋して示す。

Step1: ファーマコフォア中の特性基が検索対象の分子中に含まれているか否かを評価

```

match, mList = EmbedLib.MatchPharmacophoreToMol(mol,FeatFact,pcophore)
if match:                                     ↑ ファーマコフォアのマッチング(特性基のみ)
    num_match = len(mList)
    for i in range(num_match):                 # マッチングした特性基を構成する原子の ID を表示
        num_feature = len(mList[i])
        for j in range(num_feature):
            print mList[i][j].GetAtomIds()

```

Step2: ファーマコフォアで定義されている特性基間の距離が検索対象の分子内で満たされているか否かを評価

```

bounds = rdDistGeom.GetMoleculeBoundsMatrix(mol) ← 距離行列の取得
pList = EmbedLib.GetAllPharmacophoreMatches(mList,bounds,pcophore)
num_match = len(pList)                         ↑ ファーマコフォアのマッチング
phMatches = []
for i in range(num_match):                     # ファーマコフォアにマッチングした特性基
    num_feature = len(pList[i])               # を構成する原子の ID を取得
    phMatch = []
    for j in range(num_feature):
        phMatch.append(pList[i][j].GetAtomIds())
    phMatches.append(phMatch)                 ← phMatches は Step3 で利用する

```

RDKit では、Distance Geometry (DG) [5]により分子の 3D 構造を生成させる(実際の構造生成は Step3 で実施)。DG は原子間の距離データから分子の配座を算出する方法であり、bounds 変数に格納されているデータは、この距離データ(距離行列)に相当する。ファーマコフォアで定義されている特性基間の距離と本距離行列を比較することで、検索対象の分子内でファーマコフォアが満たされているか否かを評価している。

Step3: ファーマコフォアで定義されている特性基間の距離を分子の距離行列に埋め込み 3D 構造を生成

```

for phMatch in phMatches:
    bm,embeds,nFail                        ↓ ファーマコフォア制限付きでの 3D 構造生成
    =EmbedLib.EmbedPharmacophore(mol,phMatch,pcophore,count=20,silent=1)
    print "----> embeds num: ",len(embeds)
    for embed in embeds:
        AllChem.UFFOptimizeMolecule(embed) ← UFF を用いた最適化

```

len(embeds)が 0 でなければ、ファーマコフォアを満たす配座をもつ分子(フラグメント)であると判定することができる。また、配座生成は、20 回トライ(count=20)するように設定しているため、最大 20 個のファーマコフォアを満たす配座を得ることができる。生成された配座には、Universal Force Field による最適化を実施する。

3.5 フラグメントと結合ポケット間のマッチング評価

本稿では、CDK2 の結合ポケットに位置する阻害剤の 3D 座標からファーマコフォアを定義しているため、このファーマコフォアとヒットしたフラグメントとをアライメントすることにより、CDK2 とフラグメントの複合体構造を構築することができる。アライメントを行うための回転、平行移動行列の取得と、アライメントは以下の処理で行うことができる。

```

align_data = rdAlignment.GetAlignmentTransform(ref_mat,move_mat) ← 行列の取得
AllChem.TransformMol(mol,align_data[1]) ← アライメント

```

結合ポケットにアライメントされたフラグメントは最大 20 個の配座をもつが、それぞれの配座に対し、結合ポケットを構成する原子と接触がないか、評価を行う。本稿では、フラグメントと結合ポケットを構成する原子間の距離が 1.5 Å 以内であれば、そのフラグメントの配座は結合ポケットに接触すると判断している。一方、結合ポケットと接触しないと判定された配座を 1 つ以上もつフラグメントは、その配座情報と共に標的指向型フラグメントライブラリに格納される。以上の処理により、本創薬システムのラピッドプロトタイピングは終了となる。

3.6 プロトタイプの検証

本プロトタイプの検証を行うにあたり、A Directory of Useful Decoys (DUD) [6]に登録されている 40 種類のタンパク質の活性既知の化合物を集めて、化合物ライブラリとして用いた(2950 化合物)。次に、RECAP によるフラグメント化と Rule of Three によるフィルタリングにより、フラグメントライブラリを作成した(1512 化合物)。最後に、ファーマコフォアは、CDK2 と阻害剤の共結晶(PDB ID: 3LFN)を用いて定義し(図 3 右)、フラグメントライブラリに対するファーマコフォアサーチと CDK2 の結合ポケットとのマッチングを実施した。結果として、43 化合物からなる標的指向型フラグメントライブラリが構築された(図 4 左)。この 43 化合物は、定義したファーマコフォアを満たし、かつ CDK2 の結合ポケットにフィットしていることを確認した(図 4 右)。

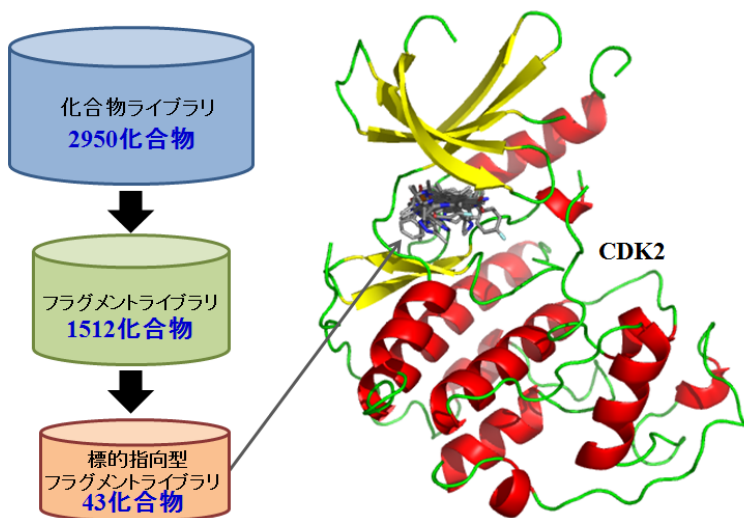


図 4 プロトタイプの検証

4. 最後に

本稿では、ラピッドプロトタイピングという視点から、Chemoinformatics Toolkits の使用法について述べた。適用例として、標的タンパク質指向型フラグメントライブラリの構築について紹介した。ソースコードの総行数は 500 行にも満たないため、API の使用法さえ理解していれば、半日もあれば書くことのできる量である。もちろん、実践的な創薬システムとしては、多くの改善すべき点はあるが、フラグメントの構造式や、標的タンパク質との結合様式は、他の研究者とのコミュニケーションの題材として、有益に利用できるものと思われる。

創薬システムの開発において、アイデアを迅速に具体化できる面白さは、Webシステム開発におけるマッシュアップに通じるものがあり、また、ToolkitsのAPIの利用法の整備は、その土台になるものと考えている。なお、RDKitのAPIに関する基本的な利用法については、サンプルプログラムも含めて、<http://cheminformist.itmol.com>で紹介している。本稿が、皆様の研究開発に少しでもお役に立てれば幸いである。

5. 参考文献

- [1] O'Boyle NM, Hutchison GR, *Chem. Cent. J.* **2008**, 2, 24.
- [2] Congreve M, et al. *Drug Discov. Today* **2003**, 8, 876.
- [3] Lewell XQ, et al. *J. Chem. Inf. Comput. Sci.* **1998**, 38, 511.
- [4] Lesuisse D, et al. *Bioorg. Med. Chem. Lett.* **2010**, 20, 1985.
- [5] Blaney JM, Dixon JS, *Reviews in Computational Chemistry*, Vol. 5, VCH:New York, **1994**, p. 299.
- [6] Huang N, et al. *J. Med. Chem.* **2006**, 49, 6789.

//// Activities ////

構造活性フォーラム 2010 開催報告
「QSAR パラダイムの分化と深化」構造活性フォーラム 2010
実行委員長 中川好秋

平成 22 年 6 月 18 日 (金), コーポイン京都 (京都市中京区) において, 構造活性フォーラム 2010 を「QSAR パラダイムの分化と深化」というタイトルで開催致しました. 今回は, QSAR の原点に立ち戻り, Classical QSAR (Hansch-Fujita 法) に焦点を当て, その生みの親である藤田先生に QSAR 誕生の秘話と, その応用例について講演をしていただきました. 今からちょうど 25 年前, 定期的なフォーラムが開催される以前の 1985 年 (昭和 60 年) 6 月に, 「定量的構造活性相関の基礎と応用」というタイトルで, 講習会が開催されました. この頃は, まだ日本薬学会の部会としては活動しておらず, 構造活性相関懇話会の主催で 2 日間にわたって行われました. 今回の講演要旨に藤田先生が書いておられますが, QSAR の誕生したのは 1961 年のことで, 来年でちょうど半世紀を迎えます.

今回のフォーラムへの参加者は講演者を含めて 101 名でありました. また, 学生の参加費を無料にしましたところ, 21 名の学生が参加してくれました. 一般参加者としては, 大学や公的研究機関から 24 名, 企業からは 56 名でありました. 当日のプログラムは以下の通りでした.

- | | |
|--|----------|
| 1. 清水良 (田辺三菱製薬) | 座長 中山 章 |
| QSAR---統計的有意性とメカニズムの妥当性 | |
| 2. 藤田稔夫 (京都大学名誉教授) | 座長 高山千代蔵 |
| Hansch-Fujita 法の半世紀---生物物理化学的意義と創薬企業化成功例 | |
| 3. 大軽貴典 (田辺三菱製薬) | 座長 大田雅照 |
| 等級活性データーの取り扱いと 3D-QSAR | |
| 4. 中馬寛 (徳島大学) | 座長 岡島伸之 |
| Hansch-Fujita 法の分子論的解釈と分子科学計算を用いた enzymatic QSAR | |
| 5. 仲西功 (近畿大学) | 座長 藤原 巖 |
| 分子シミュレーションと構造活性相関 | |

清水博士には, 藤田先生と入念な打ち合わせを行なっていただき, Hansch-Fujita 法を正しく理解するための導入的な内容で講演していただきました. その甲斐もあって, 初心者にとって分かりやすく, 藤田先生の講演にうまく連結したと思います. 藤田先生には, QSAR という手法がどのようにして誕生したかを簡単に紹介していただき, Hansch-Fujita 法が創薬研究の現場で頻繁に使われてきたことをお話ししていただきました. 座長は藤田先生のお弟子さんの高山博士にお願いしましたが, 高山博士が藤田先生から受けた薫陶が語られるとともに, QSAR 誕生時の Hansch, Fujita 両先生の写真も公開されました. 午後は, 大軽博士に 3D-QSAR 解析と等級データーの取り扱いについて, 中馬博士と仲西博士には, 計算科学の分野からの講演をお願いしました. お二人からは, それぞれ Hansch-Fujita 法の物理化学的な解釈とタンパク質とリガンド分子間の相互作用についての講演をしていただきました.

すべての講演において, 活発な質疑応答が行われ, 途中で質問を打ち切ったにもかかわらず大幅に時間超過し, 総合討論の時間を持つことができませんでした. しかし, それぞれの座長の見事な進行によって, 和やかなうちに本講演会を閉じることができました. また, 懇親会は, 例年のような形をとらず, 講演者, 学生含むすべての皆様から会費を頂戴して開催させていただきました. 予想以上に多くの方々に参加していただき, 懇親を深めることができました.

末筆ながら, ご講演を頂いた講師の先生方, 座長の労をおとりいただいた 5 人の先生方には心から御礼申し上げます.

///// Activities /////

<会告>

第38回構造活性相関シンポジウム

日時 平成22年10月30日(土)・31日(日)
 会場 徳島大学工学部共通講義棟(5-6F)(徳島市南常三島町2-1)
 主催 日本薬学会構造活性相関部会
 共催 日本化学会 日本分析化学会 日本農芸化学会 日本農薬学会
 併催 第33回情報化学討論会
 懇親会 10月30日(土)
 連絡先 第38回構造活性相関シンポジウム事務局
 〒105-0014 東京都港区芝3-17-15 クリエイト三田 207
 Tel 03-3798-5240 Fax 03-3798-5251 E-mail sar2010@event-convention.com

第1日目(10月30日)

9:30 - 9:40 開会 中馬寛(徳島大院・薬)

9:40 - 11:45 SAR Presentation Award 応募講演(工学部 創成スタジオ)

座長：赤松美紀

KA01 分子科学計算に基づくベンゼンスルホンアミド誘導体の炭酸脱水酵素阻害機構解析と相関解析

○宗井陽平、島本和典、清水美帆、相原薫、山内香子、吉田達貞、中馬寛(徳島大院・薬)

KA02 インフルエンザノイラミニダーゼ-シアル酸誘導体複合体相互作用の非経験的フラグメント分子軌道法計算に基づく相関解析

○比多岡清司、的場弘、郡恵理、原田政隆、吉田達貞、中馬寛(徳島大院・薬)

KA03 縮環インドール構造を有する新規Eg5阻害剤の開発

○竹内智起、大石真也、渡部敏明、大野浩章、浅田直也、北浦和夫(京大院・薬)、澤田潤一、浅井章良(静岡県大院・薬)、藤井信孝(京大院・薬)

座長：藤原巖

KA04 メディシナル・エレクトロノミクス分子としてのオリゴアセチレニック芳香族化合物の疎水性

○小畑勝稔、中田英司、宇都義浩、堀均(徳島大院・STS研)

KA05 代謝物同定およびIN SILICO予測を用いたヒト代謝酵素CYP3A4とCYP2C19における基質認識機構の差異

○城谷直紀、十川萌、原田俊幸、宮川恒、平井伸博、赤松美紀(京大院・農)、生城真一、榊利之(富山県立大・工)

12:40 - 14:40 ポスターセッション(工学部 K407 教室) ※情報化学討論会ポスター会場と共通

12:40 - 13:40 奇数番発表

13:40 - 14:40 偶数番発表

14:40 - 15:30 一般講演(工学部 創成スタジオ)

座長：飯島洋

- KO01*** 量子化学計算によるタンパク質とリガンドの結合に及ぼす溶媒効果の解析
○浅田直也、北浦和夫(京大院・薬)
- KO02*** 高活性・高選択性 α_2C Adrenoceptor antagonist の創製とホモロジーモデルを用いた相互作用解析
○渡辺佳晃、神辺義剛、柏木浩孝、大竹義仁、古田佳之、山口真里奈、與語健二、吉田昌伸、須藤宏和、今川純一、古賀隆樹、大田雅照、森川一実(中外製薬株式会社)

15:30 - 16:25 特別講演 (工学部 創成スタジオ)

座長：岡島伸之

- KS01** Thermodynamic Analysis of Protein-Protein and Protein-Ligand Interactions Provides Rigorous Guidelines for Drug Development
Ernesto Freire (Johns Hopkins University, Department of Biology, USA)

16:35 - 17:30 特別講演 (工学部 創成スタジオ)

座長：船津公人

- JS01** スパコンがもたらす計算化学の革新
平尾公彦(理化学研究所)

17:30 - 18:15 招待講演 (工学部 創成スタジオ)

座長：大田雅照

- KI01** 単体及び複合体タンパク質を標的としたインシリコスクリーニング
広川貴次(産業技術総合研究所 生命情報工学研究センター)

19:00 - 懇親会 (ホテルグランドパレス徳島) ※情報化学討論会と合同

ポスターセッション (工学部 K407 教室)

- KP01** PIMT が L- β -Asp あるいは D- α -Asp を含むペプチド基質を認識するメカニズムの計算機的研究
○野地郁彦、小林佳奈(東北薬大・薬)、小田彰史(東北薬大・薬、阪大・蛋白研)、高橋央宜(東北薬大・薬)
- KP02** CYP の基質および阻害剤選択性に関する統合解析-機械学習, ドッキング, QSAR-
○岡田耕平、山本将博、吉田達貞、中馬寛(徳島大院・薬)
- KP03** ヒドロキシルラジカルによるアミノ酸残基の α -水素引き抜きについての密度汎関数計算
○高橋央宜、小林佳奈(東北薬大)、小田彰史(東北薬大、阪大・蛋白研)
- KP04** 分子内 Redox 反応を利用したフルオロアルケンジペプチドイソスター (FADI) 含有ペプチドの合成と活性評価
○辻耕平、八巻陽子、山本純、重永章(徳島大院・薬)、下東康幸(九州大院・理)、大高章(徳島大院・薬)

- KP05** 分子モデリング・分子科学計算に基づくヒトノイラミニダーゼの構造-機能解析
○原田政隆、比多岡清司、郡恵理、的場弘、北尾聡、Motiur Md. Rahman, 吉田達貞、辻大輔(徳島大院・薬)、広川貴次(産総研・CBRC)、伊藤孝司、中馬寛(徳島大院・薬)
- KP06** 非経験的分子軌道法によるフェノール水素原子のラジカル引き抜き反応の分子論的考察およびフラボノイドの構造活性相関への応用
○廣隅公治、馬島彬、吉田達貞(徳島大院・薬)、志葉優子、河合慶親、寺尾純二(徳島大院・栄養)、中馬寛(徳島大院・薬)
- KP07** マウスリゾチームにおけるアスパラギン残基の部位特異的なラセミ化に関する量子化学計算
○小林佳奈(東北薬大)、小田彰史(東北薬大、阪大蛋白研)、高橋央宜(東北薬大)
- KP08** ヒト Cathepsin A 活性に対する Arg344 置換の影響に関する実験および分子科学計算に基づく解析
○郡恵理、比多岡清司、原田政隆、的場弘、北尾聡、Motiur Md. Rahman、吉田達貞(徳島大院・薬)、門田佳人(徳島文理大・薬)、辻大輔(徳島大院・薬)、広川貴次(産総研・CBRC)、伊藤孝司、中馬寛(徳島大院・薬)
- KP09** S-15176 およびその誘導体がミトコンドリアの膜構造と機能に及ぼす作用
○川島聡(徳島大・薬、徳島大疾患ゲノム研セ)、山本武範(徳島大疾患ゲノム研セ)、堀内優加、藤原健悟(徳島大・薬、徳島大疾患ゲノム研セ)、山下菊治(徳島大院・HBS)、寺田弘(東京理大・薬)、兼松誠(徳島大・薬)、穴戸宏造(徳島大・薬)、篠原康雄(徳島大・薬、徳島大疾患ゲノム研セ)
- KP10** D- α -アスパラギン酸および L- β -アスパラギン酸の分子力学計算のためのパラメータの検討
○小田彰史(東北薬大・薬、阪大・蛋白研)、小林佳奈、高橋央宜(東北薬大・薬)
- KP11** In silico 創薬技術に基づく D-aspartate oxidase-thiolactomycin 複合体の構造解析
○中込泉、山乙教之、合田浩明、片根真澄、本間浩、広野修一(北里大・薬)
- KP12** ICA によるファルマコフォアのクラスタリング
○石川誠(日産化学)
- KP13** 新規解析手法を用いた QSAR モデルの構築
○本保洋介(阪大院・薬)、岡本晃典(阪大院・薬)、川下理日人(阪大院・薬、阪大微生物病研)、大眉佳大、栗花落昇平、伊藤雅士(阪大院・薬)、高木達也(阪大院・薬、阪大微生物病研)
- KP14** 非経験的分子軌道法計算に基づく薬物-受容体分子間相互作用における Hammett σ の電子的効果の解析
○清水美帆、馬島彬、吉田達貞、中馬寛(徳島大院・薬)
- KP15** 結合自由エネルギー計算とランダムフォレストによるプロテインキナーゼ阻害活性の予測

- 田村勇之進、宮川博夫(大正製薬)
- KP16** 1-β-O-アシルグルクロニドの構造活性相関：親電子的分解反応に対するカルボン酸 α 位置換基とキラリティの影響
馬場暁子、○吉岡忠夫(北海道薬大)
- KP17** 化学物質の反復投与毒性試験における神経毒性のデータ解析とカテゴリー化の検討
○西川智、櫻谷祐企、山田隆志、山田隼、前川昭彦、林真(製品評価技術基盤機構)
- KP18** 化学物質の反復投与毒性を対象としたカテゴリーライブラリーの開発
○櫻谷祐企、西川智、山田隆志、山田隼、前川昭彦(NITE)、林真(NITE、安評センター)
- KP19** 副作用情報に基づく薬物性肝障害惹起医薬品の分類と機械学習による安全性予測
○庄野由佳理(徳島大・薬)、足立麻美、坂本久美子(徳島大病院)、佐藤陽一、山内あい子(徳島大院・薬)
- KP20** 毒性予測分野での適用を目指した新規分類/予測手法の提案
○湯田浩太郎((株)インシリコデータ)
- KP21** 化学物質の生物濃縮性における極性官能基の影響
○池永裕、櫻谷祐企、山田隼(NITE・化学物質管理センター)
- KP22** データマイニングによる反復投与毒性データからの基本活性構造(BAS)抽出
大森紀人、○堀川裕志、森幸雄、山川真透、吉岡祐一、川崎万基子、岡田孝(関学大・理工)、林真、櫻谷祐企、阿部武丸、西川智(NITE)、広瀬明彦(国衛研)
- KP23** 自然免疫関連ペプチドと bis-QAC の柔軟性と抗菌能
○大倉一人(千葉科学大学)、篠原康雄(徳島大・薬、徳島大疾患ゲノム研セ)、堀均(徳島大院・STS 研)
- KP24** ベイジアンネットワークによる化学物質の反復投与毒性評価システムの構築
○廣島亮、山口一歩、天木辰哉、岡田孝(関学大・理工)、櫻谷祐企(NITE)
- KP25** 誘導適合を考慮したバーチャルスクリーニングによる新規 HCV NS3-4A プロテアーゼ阻害剤の探索
○高谷大輔、上條加寿恵(理研)、山下篤哉、前川伸哉、雨宮史武、坂本直哉(山梨大院・医)、脇田隆字(国立感染研)、榎本信幸(山梨大院・医)、伊藤正彦(山梨大院・医)、本間光貴(理研)、梅山秀明(理研、北里大学・薬)、松本武久(理研)、横山茂之(理研)
- KP26** アミノ酸残基の周辺環境に注目した GPCR の配列特徴解析システムの開発
○家村享明、加藤博明(豊橋技科大)
- KP27** 遺伝的アルゴリズムによってタンパク質ポケットを重ねあわせる
○菅谷昇義(ファルマデザイン)
- KP28** タンパク質構造予測のための関数とホモロジーモデリングシステムの構築
○荒井まみ(中央大院・理工)、加納和彦、寺師玄記、梅山秀明(北里大・薬)、岩館満雄(中央大・理工)

- KP29** TRNOE 測定、コンピュータリガンドドッキング計算および結合自由エネルギー計算を組み合わせたキチナーゼ B と Argifin 由来ジペプチド阻害剤の相互作用解析
○合田浩明(北里大・薬)、砂塚敏明、廣瀬友靖、井口加奈美、菅原章公、野口吉彦、斉藤佳史、山本剛、塩見和朗、大村智(北里生命科学研)、広野修一(北里大・薬)
- KP30** 多数の drug-like 化合物を用いたコンホメーション発生プログラムの評価
○高木輝文、天野倫子、田中稔祐、富本昌樹(武田薬品工業)
- KP31** 薬物類似構造探索におけるデータ探索手法の比較
○Phan Thieu Van、高橋由雅(豊橋技科大院・工)
- KP32** 肝臓への副作用をおこしやすい特徴的化学構造の抽出
○吉岡祐一、大森紀人、堀川裕志、森幸雄、山川真透、岡田孝(関学大・理工)
- KP33** Gaussian09-SMD 法を用いた LogP 予測
○松本真洋、石川俊夫(石原産業株式会社)

第 2 日目 (10 月 31 日)

8:30 - 9:20 一般講演 (工学部 創成スタジオ)

座長：小沢知永

- KO03*** インフルエンザウィルス由来金属酵素 RNA ポリメラーゼを標的とするジケト酸阻害剤の結合様式の予測と分子設計
○石川吉伸、藤井敏(静岡県大・薬)
- KO04*** 原子間相互作用の摂動解析
○小山洋平(理研・CDB)、小林徹也(東大・生研)、上田泰己(理研・CDB)

9:20 - 10:20 特別セッション (工学部 創成スタジオ)

「タンパク質ホモロジーモデリングと構造活性相関の融合」

座長：久保寺英夫

- KF01** 複合体タンパク質モデリングと生物情報学的リガンドドッキングの統合：Integrated-FAMS
○梅山秀明(北里大・薬、理研)、小松克一郎、山本理恵(北里大・薬)、高谷大輔(理研)、加納和彦(NIID Japan)、寺師玄記(北里大・薬)、岩舘満雄(中央大、理研)、竹田-志鷹真由子(北里大・薬、理研)
- KF02** 昆虫脱皮ホルモン受容体とリガンドとの相互作用
中川好秋(京大院・農)
- KF03** ニコチン性アセチルコリン受容体と殺虫剤ネオニコチノイドとの相互作用
赤松美紀(京大院・農)

10:30 - 11:15 招待講演 (工学部 創成スタジオ)

座長：山下富義

- KI02** 薬物の毒性と代謝：基質の構造から酵素の認識を読む
山添 康(東北大院・薬)

11:15 - 12:10 特別講演 (工学部 創成スタジオ)

座長 : 藤田稔夫

- KS02** Condensing Chemical Reactions to Pseudo-molecules: An Efficient Way of Reactions Mining
Alexandre Varnek (University of Strasbourg, Laboratory of Chemoinformatics, France)

13:10 - 14:40 招待講演 (工学部 創成スタジオ)

座長 : 宍戸宏造

- KI03** エピジェネティックに遺伝子発現を制御する低分子化合物の設計と合成
宮田直樹(名古屋市立大院・薬)

座長 : 大高章

- KI04** 細胞治療を助ける化合物
上杉志成(京大・物質-細胞統合システム拠点)

14:55 - 16:15 一般講演 (工学部 創成スタジオ)

座長 : 合田浩明

- KO05*** 自然言語処理による CYP 代謝情報の網羅的収集とそれに基づく構造活性相関解析
○山下富義、吉田秀哉、馮春来、橋田充(京大院・薬)

- KO06** 進化計算を用いた新薬候補構造の骨格デザインに関する研究
○丸野裕史、高橋由雅(豊橋技科大院・工)

座長 : 本間光貴

- KO07** 相互作用プロファイルによるタンパク質複合体予測のポストドッキング解析
○内古閑伸之(東工大院・情理)、広川貴次(産総研・CBRC)、秋山泰(東工大院・情理)

- KO08*** 分子重ね合わせ法を用いた Ligand-Based Drug Design 手法による congeneric な化合物群
の 3D-QSAR モデルの構築及びその検証
○土井一生、山乙教之、合田浩明、広野修一(北里大・薬)

16:15 - 16:30 SAR Presentation Award 表彰式 (工学部 創成スタジオ)

閉会

※一般講演 (KO01~08) の*印は A 講演 25 分 (発表 20 分, 質疑 5 分), 無印は B 講演 15 分 (発表 10 分, 質疑 5 分), SAR Presentation Award 応募講演 (KA01~05) は 25 分 (発表 18 分, 質疑 7 分)

//// Activities ////

構造活性相関部会の沿革と趣旨

1970年代の前半、医農薬を含む生理活性物質の活性発現の分子機構、立体構造・電子構造の計算や活性データ処理に対するコンピュータの活用など、関連分野のめざましい発展にともなって、構造活性相関と分子設計に対する新しい方法論が世界的に台頭してきた。このような情勢に呼応するとともに、研究者の交流と情報交換、研究発表と方法論の普及の場を提供することを目的に設立されたのが本部会の前身の構造活性相関懇話会である。1975年5月京都において第1回の「懇話会」(シンポジウム)が旗揚げされ、1980年からは年1回の「構造活性相関シンポジウム」が関係諸学会の共催の下で定期的開催されるようになった。

1993年より同シンポジウムは日本薬学会医薬化学部会の主催の下、関係学会の共催を得て行なわれることとなった。構造活性相関懇話会は1995年にその名称を同研究会に改め、シンポジウム開催の実務担当グループとしての役割を果たすこととなった。2002年4月からは、日本薬学会の傘下組織の構造活性相関部会として再出発し、関連諸学会と密接な連携を保ちつつ、生理活性物質の構造活性相関に関する学術・研究の振興と推進に向けて活動している。現在それぞれ年1回のシンポジウムとフォーラムを開催するとともに、部会誌のSAR Newsを年2回発行し、関係領域の最新の情勢に関する啓蒙と広報活動を行っている。

本部会の沿革と趣旨および最新の動向などの詳細に関してはホームページを参照頂きたい。

(http://bukai.pharm.or.jp/bukai_kozo/index.html)

編集後記

日本薬学会構造活性相関部会誌SAR News第19号をお届けいたします。今回は、創薬におけるchemoinformatics手法の活用という観点から纏めてみました。お忙しい中ご執筆いただきました諸先生方に、心よりお礼申し上げます。Perspective/Retrospectiveでは、長谷川清先生(中外製薬)に、ファーマコフォアモデルによるバーチャルスクリーニングやde novo設計などchemoinformatics分野の最先端研究・最新情報のご紹介とともに、プロジェクト化合物最適化のための独自ADMET予測システム構築や化合物空間の可視化を解説していただきました。Cutting Edgeでは、大川和史先生(持田製薬)に、化学構造プログラミング・化合物構造最適化プロジェクトシミュレーション・メタ化合物管理システムなど、独創的手法の開発と活用について、また、吉森篤史先生(理論創薬研究所)に、ソフトウェアの高効率開発の実例として、オープンソースchemoinformatics toolkitsを駆使したラピッドプロトタイピング手法について、ご紹介をいただきました。創薬研究者としては、日々online/offlineで発信される膨大なchemoinformatics情報とこれを解析する新規な手法から目を離すことができません。このような状況の中、創薬現場で適時・適切にchemoinformatics手法を活用する上で、いずれのご解説も実践的・具体的情報としてたいへん有用であると思われます。このSAR Newsが構造活性相関研究の先端情報と展望を会員の皆様にご提供できることを、編集委員一同願っております。(編集委員会)

SAR News No.19 平成22年10月1日

発行：日本薬学会 構造活性相関部会長 赤松 美紀

SAR News 編集委員会

(委員長) 久保寺 英夫

藤原 巖

福島 千晶

粕谷 敦

飯島 洋

* 本誌の全ての記事、図表等の無断複写・転載を禁じます。