

医薬品および化学物質安全性研究を支援する基盤システム

国立医薬品食品衛生研究所 化学物質情報部
神沼二真[#]・中田琴子・中野達也・高井貴子

Development of Research Support Environment for Molecular Pharmacology and Toxicology

Tsuguchika Kaminuma[#], Kotoko Nakata, Tatsuya Nakano and Takako Takai

The Chem-Bio Informatics division has developed the institute-wide computer network which is connected to the Internet via IMnet (Inter Ministry Network) sponsored by the Government. We call the network and associated information and computing infrastructure as NICI, the NIHS Information and Computing Infrastructure. The primary goal of NICI is to provide general communication environment such as e-mails, and the second goal is to provide WWW environment. Accomplishing these two goals the division is now challenging to develop information and computing environment that directly supports pharmaceutical and toxicological research. The prototype of this environment which consists of various chemical and biomolecular databases and computing tools has been developed under the name of MOBI-DICS. The concept and vision on how to extend the prototype to institute-wide infrastructure was discussed.

Keywords: WWW, MOBI-DICS, computational chemistry, bioinformatics

1. NICIの将来

化学物質情報部は、平成1年度より全所的な情報基盤の構築を部の新たなミッションとして掲げ、継続的な努力を続けてきた。最初はイーサネットでは数台のワークステーションとPCを結んだ簡単な所内ネットワーク (LAN) の実験からスタートしたが、平成5年度には、インターネットに接続する環境を構築し、平成7年度には本格的なネットワークを整備し、さらに平成11年10月と12年4月にはこれを拡充した¹⁾。インターネットの利用のための基幹環境、情報コンテンツの提供環境、各種の計算リソースから構成されているこの情報計算基盤を、現在はNICIと呼んでいる。NICIとは、NIHS Information and Computing Infrastructureを意味する。

NICIの現状については別に報告があるのでここでは省略するが、情報部ではNICIの次の発展段階として、所の基幹的な研究業務である医薬品や(食品を含む)化学物質の安全性に関わる研究を支援する基盤環境の構築を開始した。現在はそのプロトタイプであるMOBI-DICSを開発している段階である。

MOBI-DICS (Molecular Biodynamics Simulator) とは、生

体には異物である化合物 (Xenobiotic Chemicals) と生体系との相互作用を分子レベルで解析し、予測するための基盤システムである。もちろんこれは単一のシステムではなく、いくつかのデータベースや知識ベース、計算ツール、情報や知識のコンテンツをインターネットとWWW技術でゆるく結合した複合システムである。それゆえMOBI-DICSは、外部の同様なシステムと容易に連結され、さらに大きな複合系を構成できる可能性を秘めている。このことは、MOBI-DICSがわれわれの部だけの研究の道具ではなく、医薬品の効率や安全性、化学物質の安全性に関わっている、所内所外の多くの研究者の支援システムとなりうることを意味している。

以下では、MOBI-DICSの概要をその開発思想と今後の計画を含めて紹介し、全所的な研究支援基盤に拡張する具体的な手順を明らかにする。

2. MOBI-DICSの概念

2.1 個々の化合物のデータベース

MOBI-DICSの開発思想は単純である。第一の目標は、相互作用の主役である生体外からの化合物 (Xenobiotic Chemicals) と生体系の構成分子それぞれのカタログをまず作成し、次にそれらの間を関係づけることである。

このうち最初に着手したのは、国際がん研究センター (IARC) が刊行しているモノグラフに準拠した発がん物質に関するデータベースである。つぎに科学技術庁の振興調

[#] To whom correspondence should be addressed:

Tsuguchika Kaminuma; Kamiyoga 1-18-1, Setagaya, Tokyo 158-8501, Japan; Tel: 03-3700-9540 Fax: 03-3700-7592; e-mail: kaminuma@nihs.go.jp

整費による研究として、内分泌かく乱物質を含む各種の生体影響データベースを開発した²⁾。現在は、わが国で申請認可されているすべての医薬品の識別情報のデータベースを開発中である³⁾。これらの化合物データベースの特徴は、X線解析や理論計算で求めた原子の三次元座標データが含まれていることである。その理由は分子の三次元構造は構造活性相関研究の基盤となると考えているからである。

生体分子に関しては、核酸の塩基配列データ、蛋白質のアミノ酸データ、三次元構造データなど一般に公開されている情報源にアクセスできる環境が整えられている。

2.2 受容体

相互作用の最もわかりやすい例は、薬物分子とその受容体との結合である。われわれは受容体を含めた Receptor Database (RDB) を開発している⁴⁾。このデータベースは生体分子のうち、受容体と考えられるものを探して、データベースとしたものであるが、当然それぞれのリガンドについての情報も入れている。

しかし、医薬品でさえそうであるが、生体に作用する化合物(生体影響物質)ではあっても、標的分子が判明していないものはまだ沢山ある。反対に受容体ではあっても、生体内物質としてのリガンドが判明していないもの(オーファン受容体)や、どのような生体外化合物が結合するかわからないものも多い。すなわち生体系分子への「リンクの張れない」化合物も少なくないのである。だが生物医学の進歩は急速であり、双方の間のリンクは、実際には急速に複雑なものになってきている。問題はそうしたリンクの増大、すなわちデータや知識の増大に追従し、現実のリンクをどのように増やしていけるかである。本データベースシステムではシステム内でデータ編集を行っており、インターネット上の遺伝情報データベースや解析システムとリンクして効率よく様々な情報を提供することに改良を重ねている。

2.3 結合実験のデータベースと構造活性相関解析

開発の第二の目標は、上記の関係をダイナミックスのモデル系として表現することである。リガンドとなる化合物 a と、受容体 A との関係が知られている場合、問題は a と A の関係だけでなく、 a の類縁物質 a' , a'' ……と A のファミリー(分子種) A' , A'' ……との関係に発展する。例えば、われわれは性ホルモンあるいはその類縁物質と受容体、例えばエストロゲン受容体との結合実験の報告をデータベース⁵⁾ にしている。Binding Affinity に注目したこのようなデータベースは、そのまま構造活性相関研究の基礎になる。

例えば、 a' , a'' ……に共通する構造上の特徴を求める問題がある。これを空間的な広がりとするれば、pharmacophore を求める問題となる。 A' , A'' ……に共通する(ホモログスな)配列や立体構造上の特徴、作用部位の同定などは、受

容体モデリングの問題となる。さらに a 群と A 群との結合ダイナミックスの解析や予測は docking study と呼ばれている。実験で得られる結合の強さや順位を計算で再現するのはまだ極めて難しい。したがって化合物と受容体のそれぞれに随伴するデータや知識をできるだけデータベースに収録しておくべきであろうと考えている。

2.4 分子の働きの記述

開発の第三の目標は、化合物の作用の相手である生体分子の生体内での存在部位と働きを記述する方法を開発することである。薬物作用の場合は、小腸における輸送、肝臓における代謝、標的臓器における主作用などに分けられる。さらにそれぞれの組織においては、標的となる酵素や受容体が細胞のどの部位に存在し、どのような働きをしているかが問題となる。そうした働きは細胞内の代謝、細胞信号伝達、細胞周期の制御、細胞死などと関係している。すなわち生体内の分子の機能は、細胞の内外における分子間の相互作用のネットワークによって発揮される。われわれは細胞信号伝達系をグラフとして記述するデータベース(CSNDB)⁶⁾ を開発してきた。さらにこれをドイツのグループが開発している転写因子のデータベース(TRANSFAC)と結合することを試みた。最近、生体内分子の相互作用ネットワークの広がりを pathway として、記述する試みが盛んになってきた。こうした記述法は現実を極端に簡素化したモデルであるが、より精密な記述への出発点として重要である。

このようなモデルを使えば、毒性化合物や、薬物に対する細胞レベルの応答を記述することができる。しかし、これらは例えてみれば、道路網のような静的なモデルであり、交通状態を記述できるダイナミカルなモデルではない。さらに、たとえ細胞レベルの反応が記述できたとしても、それが生体全体の応答とどのように関係しているのかを教えてはくれない。それにはより高次の記述が必要である。

2.5 Gate Point から Endpoints に到る因果の連鎖

われわれの最終目標は、生体外の化合物あるいはその代謝物と生体内の標的分子との結合(Gate Point)から生体(あるいは実験系)の究極の応答(Endpoints)に到る因果の連鎖を分子間の相互作用や、細胞間の相互作用として記述する方法の開発である。われわれは、これまで第一から第三の目標に挑み、プロトタイプとなるシステムを開発してきた。現在は、これまで開発してきた要素システムを薬物の研究や、毒性研究の実際問題に適用しながら第四の目標にどう挑んだらよいかを模索している。具体的には内分泌かく乱物質問題を例として、システムの検討を始めている⁷⁾。

3. 具体的な開発計画

MOBI-DICSの個々の要素システムについては⁸⁾内容を更新中の医薬品データベースを除きすでに公開使用実験を重ね、改良を続けている。MOBI-DICS実現の一番大きなポイントは、要素システムを如何に連結するかである。幸い、MOBI-DICSのすべての要素システムには、化合物(低分子化合物および生体高分子化合物)が含まれている。さらに、それらの多くにはCAS番号がつけられている。

そこで、各データベースとそこに含まれている化合物のリスト(インデックス)をまず作成し、これを手掛かりに、すべてのデータベース同士を連結することが可能である。ただし、この方法では低分子化合物データベース同士、生体高分子化合物データベース同士は統合できるが、双方を連結することはできない。そのためには、低分子化合物の中に標的となる酵素や受容体など生体高分子の情報があり、逆に生体高分子化合物のデータベースの中に低分子化合物の標的となるリガンドや結合に関する情報がなければならない。残念ながら、医薬品においてさえこうした情報は完備していない。したがって、われわれ自身で医薬品や毒性化合物ごとに、それらの標的となる生体分子をしらべていかなければならない。これは手間のかかる仕事であるが、現在少しずつ進めている。

こうした手法をベースにした連結のための作業は現在進んでおり、2000年の夏までにWWW対応のプロトタイプを完成できる予定である⁹⁾。

4. 全所的な研究支援基盤システムへの拡張

われわれの究極の目標は、医薬品開発や毒性研究を全体として支援する基盤システムやツールの開発であり、MOBI-DICSはそのためのプロトタイプである。したがって、問題はそれをどのように全所的な基盤システムに拡張し、国立衛研のCOEとしての機能を担う開かれた基盤的なリソースに発展させていくかである。

4.1 データベースの拡充

MOBI-DICSに含まれている化合物データベースはわれわれが長い時間をかけて、少しずつ整備してきたものがほとんどである。だが世界の同じような化合物データベースと較べると数は少ない。そこでMOBI-DICSの第一の拡充は、フリーで手に入る化合物データベース(例えば米国立がん研究センター(NCI)の約127,000化合物の分子構造データベース)や商品となっている化合物データベースを購入するか、ネットワークなどで閲覧可能な環境をつくることである。

例えばMDL社から購入している化合物データベースとして、ACD-3D, CMC-3D, MDDR-3Dなどがある。これらを

所内のユーザに利用できるようにすることは可能である。ただしこれらはインターネットでアクセスしてくる外部ユーザに公開することはできない。

4.2 医薬品および候補化合物ライブラリの開発

わが国で使われている医薬品は一般名では数千、商品名では数万のオーダーである。先進各国の医薬品の数もだいたい同じくらいのオーダーである。一方、医薬品会社がスクリーニング対象とする化合物は100万のオーダーであり、化合物プロバイダーと呼ばれる大手企業が扱っている化合物の数も数十万から数百万のオーダーである。現在、情報部は有機化学部と協力して、JANのデジタル化に取り組んでおり、平成12年の夏までに一般名の入力終了する予定である。これ以後は年に数十品程度のペースで更新していけばよいので管理可能と考える。

医薬品研究や毒性の基盤となる化合物データベースは、他の部でも開発されている。例えば生薬部はわが国の植物と成分化合物のデータベースを開発している。変異遺伝部はスクリーニング実験データを有しており、報告集も出しているから、それをデータベースとするのは簡単である。食品や食品添加物データなども、それぞれ独自に化合物データベースにでき、それらをMOBI-DICSと連結することもできる。内部のユーザに限定すればこれらのデータベースは統合的に利用できるし、共同研究者の利用に呈することも可能であろう。

このような国立衛研の化合物のデータベースは全体として生体に何らかの作用がある化合物のデータベースとして、わが国の研究者の基本レファランスになるであろう。例えば米国立がん研究センターの分子構造データベースNCI-3D(<http://chem.sis.nlm.nih.gov/>)やNIEHSのNTP(<http://ntp-server.niehs.nih.gov/>)などの公開データベースはNCIやNIEHSのステイタスを高からしめるものである。国立衛研でもそうした基盤的な化合物データベースを開発して公開すべきであろう。それは国立衛研をCOEたらしめる助けとなるだろう。

4.3 薬物代謝、輸送の知識ベース

われわれはCBI学会との協力プロジェクトとして、PHII計画を提唱しているが、その中でチトクロームP450(CYP)など薬物代謝酵素やトランスポータに関わるデータや知識を整理してデータベース化しようとしている。これについてはすでに専用のウェブサイトを開設しているが⁹⁾、とくにCYPについては、データベースに基づき、質問に応じて副作用や相互作用を警告するプロトタイプシステムの形で開発している。

ただし、われわれの知識ベースは米国の医薬品に関する情報を基礎にしているため、日本の医薬品に関する情報が必要である。こうしたデータや知識は頻繁に更新されねば

ならない。したがって情報コンテンツの整備は、専門家のチームで行われるべきであると考えている。

4.4 Receptor DBの拡充

現在の受容体データベース (RDB) には、受容体タンパク質の構造および機能に関する情報、遺伝情報、結合親和力情報、転写因子や転写領域に関する情報等が含まれているが、これらの内容をさらに充実する。さらに各国の医薬品データベースと相互に連携させるとともに、連携の基礎となるリガンド情報を拡充することを心がけている。

4.5 CSNDBの拡充

CSNDBにはさまざまな拡張方向が考えられる。第1はヒトに限定されている現在の知識を充実することである。第2は多細胞生物、とくに分子生物学のモデル生物とされる各種の生物のシグナル伝達系の知識を加えることである。第3は、医薬品開発や臨床応用を考慮して、目的指向にパスウェイやネットワーク分類して、知識を充実していく方向である。さらにより柔軟な経路検索機能を持たせること、各ノードにおける分子の相互作用の記述をより詳しくすることなどである。いずれも、これまでのような小人数では開発を続けられないので、体制づくりも急務である。

4.6 解析、計算ツールの拡充

各種の分子計算技法や配列解析のプログラム、結合解析、構造活性相関解析、構造からの毒性予測など、さまざまなソフトウェア (ツール) が必要である。これらについては、フリーソフトを導入したり、あるいは商品ソフトを購入したりして整備している。その中のあるものはNICIの共有リソースとして整備されたものであるが、情報部独自で整備したものもある。とくに、フラグメント分子軌道(MO)法は、中野らが共同研究者の北浦らと開発している生体高分子系も扱えるMO法の高精度の近似計算方法である⁽⁵⁾。いずれにしても、これらを活用しなければ宝の持ち腐れである。これからはリソースの整備状況を知らせて、利用者を募るなどして、この種のリソースのユーザを増し、利用スキルの向上にも務めたいと考えている。以下にこのようなリソースをあげる。

非経験的分子軌道法プログラム	Gaussian98
分子設計支援ソフトウェア	InsightII
配列解析プログラム	GCG
3D構造活性相関プログラム	CoMFA
フラグメント分子軌道法プログラム	ABINIT-MP

5. 応用事例

MOBI-DICSとその拡張システムは、国立衛研のさまざま

な研究業務の役に立つのではないかと考えている。また研究に関係した行政対応にも役立つと考える。以下にいくつかの例を挙げる。

5.1 SNPsプロジェクトへの対応

欧米のSNPsコンソシアムを嚆矢として、わが国でもSNPs探しのプロジェクトが続々と旗揚げし、データの蓄積も急ピッチで進んでいる。SNPsはマイクロサテライトと共に、塩基配列の典型的なバリエーションである。これらのデータは疾病に関連した遺伝子の探索などに利用される。ただし、MOBI-DICSへの応用はより限定的であり、薬物代謝酵素やトランスポータ、受容体やCSNDB、転写因子データベースに集録されているタンパク質のバリエーションのデータとなる。そうしたデータは人種差や個人差に対応したものである。この視点で見れば、SNPsデータが増えるごとに、その中から、MOBI-DICSの要素データベースの中のタンパク質と関係しているものだけを選択的に取り出して、該当するタンパク質の情報として追加しておくことが必要である。これは、SNPsデータ源とある種のデータベースの中のタンパク質とを結ぶ問題となる。この種の問題を解決するソフトウェアが極めて一般的な性格を有していることは明らかであろう。われわれはすでにこうしたプログラムを開発中である。

5.2 マイクロアレイ、プロテオミックスデータの解析

マイクロアレイやプロテオミックスは、ゲノムワイドな遺伝子発現やタンパク質合成を計測する革新的な技術として注目されている。しかし、これらの計測方法で生成される大量のデータをどのように解析するか、解釈するかが新たな問題として浮上してきている。これに関しては、Silicon Geneticsなど商品となっているパッケージソフトウェアがいくつも開発されているが、まだ不十分である。われわれは、薬物代謝酵素関連知識ベースやCSNDBがこうした問題に応用できると考えている。これについてはCBI学会のグランドチャレンジの一課題であるMicro AI計画として共同研究者を募って具体的な開発を進めるべく準備中である。

5.3 リスクアセスメントの方法論

MOBI-DICSは化学物質のリスク評価研究の基盤でもある。まずMOBI-DICSには、リスク評価が行われている化合物情報を、ファクトデータとしてデータベースに直接収録するか、あるいはそうしたデータベースへリンクしている。これによりユーザはリスク評価がなされている化合物ごとの情報が入手できるようになっている。また異なる生物を用いた実験データを比較するための便利な環境も用意されている。さらに現在、例えばIPCSのリスク評価の方法論などを参考として、われわれのいうGate Pointから

Endpointsに到る因果の連鎖を知識ベースとするような研究を始めている。

ただし、近未来に最も確実に期待できるのはDNAチップを用いた毒性の判定システムの開発だろう。こうした開発にはDNAチップからの大量データの解析が不可避であり、薬物代謝酵素やトランスポータに関する知識ベースやCSNDBはそうした解析に重要な働きをするだろうからである。

5.4 行政対応業務の支援

(1) QSAR

MOBI-DICSは行政からの依頼による研究にも対応している。例えばQSAR（定量的構造活性相関）である。これは化合物の環境影響をコンピュータプログラムで判定させられないかという可能性であり、OECDの化学物質安全性会議などで、主として米欧日の3者で検討されてきた。残念ながらわが国にはこの問題で行政をバックアップできる研究者がいなかった。われわれはMOBI-DICSを開発整備しながら、この要請に応える研究も始めている。

(2) 基盤研構想

大阪の基盤研究所（仮称）構想は、少しずつ論議が深まっている。最終案がまとまるまでには、まだ紆余曲折があると思うが、われわれも多少の案を提言している。いずれにしても、もし基盤研が創薬を中心とする研究所になるようならば、国立衛研とは研究面においても強い絆で結ばれるであろう。このときMOBI-DICSはまさに2つの研究所を結ぶ絆として機能するだろう。

6. 外部との協力関係

MOBI-DICSと関連リソースは所内利用や厚生省からの行政対応業務だけでなく、国内、国外の研究者に広く利用してもらうことが理想である。実際MOBI-DICSのリソースの一部はすでに国際的な評価を得ている。また、JANの電子化をめざすわが国で認可されている医薬品に関するデータベースは、将来、医薬品関係者の必須の情報源になるであろう。また薬物代謝酵素やトランスポータの知識をベースとした副作用や相互作用警告システムは、わが国の医療の薬剤部門で必要不可欠の情報リソースとなるであろう。すなわち、MOBI-DICSとその拡張システムは、医薬品や安全性研究の研究者だけでなく、臨床においても利用されうる重要なリソースとなるであろう。

このようにMOBI-DICSとその関連システムが広く使われるようになれば、ユーザからのフィードバックにより、システムを改善していくことができる。それ以上に興味深いのは、ここに一種の研究者コミュニティが形成される可能性があることである。MOBI-DICSはこのようなコミュニティによって育てられるべきであると考え。言い換え

れば、これはMOBI-DICSを基盤リソースとするインターネット上の仮想研究所を意味する。こうした仮想研究所を産官学の共同研究のネットワークと見ることもできる。インターネット時代のCOEとは、こうした仮想研究所を組織できる能力のある研究センターでなければならない。これは国の研究機関としてのミッションでもある。

MOBI-DICSの要素システムは、これまでも多くの外部の研究者や協力者の協力で開発されてきた。そうした協力者はCBI学会の活動の中で得られたものである。NICIの基礎づくりは、CBI研究会の人的なネットワークに支えられていた。MOBI-DICSが順調に発展していけるかどうかは、これからもCBI学会¹¹⁾を介した協力関係がうまく行くか否かにかかっていると言える。

7. おわりに

われわれは国立衛研の情報計算基盤であるNICIを段階的に開発、整備してきた。その第1段階は全員の電子メールの利用をめざしたコミュニケーションのためのネットワークであるが、これはいまや研究所に限らず、すべての組織に必要なインフラストラクチャーとなった。第2段階はWWWによる情報の収集と発信のための環境であり、検索エンジンやデータベースの統合検索環境である。これは研究所の研究効率の向上と共に、研究成果の社会への還元、情報公開や説明義務（Accountability）に対応したものである。この2つの段階については、すでに開発と公開実験が終わり、安定的な成長路線に移っている。

現在取り組んでいる第3段階は、研究業務そのものの支援である。そのプロトタイプが開発中のMOBI-DICSである。こうしたシステムは医薬品研究や化学物質の毒性に関わる研究など、当所の幅広い研究を支援する基盤環境である。それはまた、国立衛研をCOEたらしめる戦略的な環境である。

上記の第1と第2段階の開発により、情報部がこの所にとって、また厚生省にとって、「役に立つ」研究グループであることは証明できたと考えている。第3段階は情報部が国立衛研のある分野でのCOEたらしめるとともに、自らも優れた研究グループであることを証明するための挑戦である。しかし、これまでと同じように今度の挑戦課題も、他の部の協力を必要とする。この新しい挑戦に他の部の協力が得られることを願っている。

謝 辞

MOBI-DICSの要素データベースの開発には、山本都、大竹千代子、山本美智子、長谷川式子、瀧明子、小峰啓氏らの協力をえている。またシステム開発については、アドイン研究所の協力をえている。ここに感謝する。

参考文献

- (1) 中田琴子, 中野達也, 高井貴子, 小峰 啓, 神沼二真: 国立医薬品食品衛生研究所における研究情報基盤整備の現状. 国立医薬品食品衛生研究所報告(本号)
- (2) 神沼二真, 中田琴子, 中野達也, 高井貴子: 内分泌攪乱物質(環境ホルモン)研究を支援する情報計算基盤の概念について, 第21回情報化学討論会講演要旨集 pp90-93 (1998.11 東京)
- (3) 鈴木聡子, 川出達, 瀧明子, 長谷川式子, 山本美智子, 中野達也, 宮田直樹, 神沼二真: 日本医薬品一般名称のデータベース化, 第1回CBI学会(2000.7 東京)
- (4) Nakata, K., Takai, T. and Kaminuma, T.: Development of A Receptor Database: Receptor DB. *Bioinformatics*, vol.15, pp.544-552 (1999)
- (5) Kaminuma, T., Takai-Igarashi, T., Nakano, T., and Nakata, K.: Modeling of signaling pathways for endocrine disruptors, *Bio Systems* 55 (2000) pp.21-31
- (6) Takai-Igarashi, T. and Kaminuma, T.: A Pathway Finding System for the Cell Signaling Networks Database, *In Silico Biology*, 1, pp.129-146, (1999).
- (7) MOBI-DICS, 受容体データベース, CSNDBのURLは以下の通りである.
<http://dcbint3.nihs.go.jp/MOBI/>
<http://impact.nihs.go.jp/RDB.html>
<http://geo.nihs.go.jp/csndb>
- (8) MOBI-DICSの統合, *BIO Systems* 55 (2000)
- (9) PHIIのURL: <http://molddb.nihs.go.jp/PHII/>
- (10) Kitaura, K., Sawai, T., Asada, T., Nakano, T. and Uebayasi, M.: *Chem. Phys. Letters* 312 (1999) 319-324; Kitaura, K., Ikeo, E., Asada, T., Nakano, T. and Uebayasi, M.: *Chem. Phys. Letters* 313 (1999) 701-706; Nakano, T., Kaminuma, T., Sato, T., Akiyama, Y., Uebayasi, M. and Kitaura, K.; *Chem. Phys. Letters*, 318 (2000) 614-618
- (11) CBI学会のURL: <http://www.cbi.or.jp/>