

# Laboratorio di Preparazioni Estrattive

Introduzione alla Estrattiva



SAPIENZA  
UNIVERSITÀ DI ROMA

## Premessa

Negli ultimi anni sono incredibilmente aumentati sia il numero che la sofisticazione degli strumenti a disposizione del chimico farmaceutico per la progettazione, identificazione e studio di nuovi farmaci. Tra questi metodi, possiamo citare la chimica combinatoriale (generazione di un gran numero di molecole), la chimica computazionale (simulazioni), le biotecnologie ed i metodi di high-throughput-screening (HTS) (test di 100-1000 composti al giorno).

## La chimica combinatoriale e lo screening ad alta capacità

L'evoluzione scientifica della ricerca farmaceutica è stata costante nel corso degli anni, ed ha significato un continuo beneficio per la popolazione dei pazienti nella cura di varie importanti affezioni. Purtroppo l'evoluzione contemporanea del panorama industriale farmaceutico mondiale ed italiano ha generato conseguenze negative per le compagnie farmaceutiche. I due fenomeni di maggior impatto sono:

- l'aumento dei costi per lo sviluppo di nuovi farmaci sul mercato, cresciuti di otto volte dal 1976 (circa 50M di dollari) al 1998 (circa 400M di dollari);
- l'aumento dei tempi necessari per lo sviluppo di un nuovo farmaco, raddoppiati dal 1960 (circa 7 anni) al 1998 (circa 14 - 15 anni).

## La chimica combinatoriale e lo screening ad alta capacità

L'aumento dei costi e la riduzione del tempo nel quale sfruttare il potenziale economico di un farmaco prima del termine della copertura brevettuale ha dato un nuovo impulso verso la ricerca di processi meno costosi e/o più rapidi per la scoperta e lo sviluppo di un nuovo farmaco.

L'approccio farmaceutico classico, che è ancora utilizzato ampiamente, consiste nella sintesi di un prodotto usando metodi sintetici tradizionali; nella sua purificazione e caratterizzazione strutturale; nella determinazione delle sue proprietà biologiche sul target analizzato usando saggi biologici convenzionali; più in generale, nel verificare la sua aderenza a un profilo ideale del farmaco. L'affidabilità di questo approccio classico è molto alta, ed infatti nel corso degli anni molte molecole sono state scoperte ed immesse sul mercato. Sfortunatamente, la capacità o throughput sia della sintesi che della caratterizzazione biologica è estremamente ridotta; in prima battuta, la sintesi e la caratterizzazione di 100 molecole complesse per anno può essere considerata un successo per un piccolo gruppo di ricercatori.

# La chimica combinatoriale e lo screening ad alta capacità

Il primo "scossone" al paradigma classico della ricerca farmaceutica è stato causato dai cosiddetti saggi biologici ad alta capacità, o high-throughput screening (HTS) assays. Questi saggi, mirati a determinare l'attività di composti su un target molecolare, richiedono quantità minime di composto da testare e di reagenti e sono facilmente automatizzabili. La loro introduzione, iniziata all'avvio degli anni '80 ed ormai consolidata in tutte le grandi compagnie farmaceutiche, ha causato un enorme aumento della "domanda" di composti chimici da testare: un saggio HTS può facilmente analizzare decine di migliaia di composti per settimana, ma se solo pochi prodotti sono disponibili la produttività totale del processo di scoperta del farmaco (Drug Discovery) non è aumentata.

# La chimica combinatoriale e lo screening ad alta capacità

All'inizio le compagnie si sono orientate verso lo screening sistematico delle collezioni di composti generate durante molti anni di attività (tipicamente centinaia di migliaia di prodotti), usando il cosiddetto "random screening approach"; queste collezioni sono state ripetutamente testate su ogni target molecolare di interesse per ogni azienda, ed hanno permesso di identificare varie molecole ad attività "non prevedibile razionalmente". Questo approccio è tuttora molto popolare nelle grandi aziende farmaceutiche, e viene sempre più mediato da strumenti computazionali per definire la similarità e/o la diversità delle collezioni di composti sottoposte ad HTS.

## La chimica combinatoriale e lo screening ad alta capacità

Queste collezioni, seppur utili, non sono certo ideali come fonti di nuovi principi farmaceutici; infatti, esse sono fortemente influenzate dall'attività storica di un gruppo e contengono molte strutture molto simili; in una parola, non sono molto "diverse" e non permettono di esplorare adeguatamente il cosiddetto "spazio chimico" dal quale pescare nuovi ed originali principi attivi. A questo proposito, sono oggi disponibili metodi computazionali che permettono di "disegnare" molecole singole o addirittura set numerosi di molecole (chiamate "librerie virtuali") allo scopo di diversificare un set per HTS; diventa quindi sempre più importante avere a disposizione metodi sintetici per produrre simultaneamente e rapidamente un gran numero di composti strutturalmente correlati con buona purezza. L'insieme di queste tecniche chimiche ad alta capacità (high-throughput synthesis) è comunemente chiamato chimica combinatoriale

## Premessa

Il ruolo del chimico farmaceutico di sintesi è messo sempre più alla prova specialmente dalla notevole mole d'informazioni derivanti dalla conoscenza sia delle strutture che delle funzioni di bio-macromolecole come enzimi, recettori, trasportatori di membrana, canali ionici e dalle accresciute nozioni in settori come la genetica, la biologia e la patologia.



# Bio·IT World

INFORMATION TECHNOLOGY FOR THE LIFE SCIENCES

October 2002 • VOL. 1, NO. 8

## Banking on Structures



An explosion of structural information is on the horizon,

and the Protein Data Bank — the single international repository for data

on the three-dimensional structures of biomolecules — is ready.

BY TRACY SMITH SCHMIDT

Except for the small sign at the end of the walkway, you wouldn't guess that the nondescript, low-level brick building on the campus of Rutgers University is home to one of the world's most important biological databases — the Protein Data Bank (PDB). It seems too insignificant to contain all the personnel and equipment needed to manage the current flow of information into and out of the database, let alone handle the approaching surge of structural data.

Indeed, it is too small. The Rutgers location in Piscataway, N.J., is only one of three sites responsible for the PDB, along with the San Diego Supercomputer Center (SDSC) at the University of California, and the National Institute of Standards and Technology (NIST) in Gaithersburg, Md. Collectively, they make up the Research Collaboratory for Structural Bioinformatics (RCSB). Rutgers is responsible for data processing, SDSC for managing the public database, and NIST for maintaining the physical archive of PDB files. Together, they appear well-equipped to handle as much structural data as the community can throw at them.

The PDB harbors the coordinate files that pinpoint the location of nearly every atom in thousands of proteins and nucleic acids — maps obtained using structural techniques such as X-ray crystallography and NMR (nuclear magnetic resonance) spectroscopy (see "Solving Structures," page 3).

Imaging programs can interpret these files, allowing scientists to visualize the convoluted shape of a protein and assess how it may interact with carcinogens or drugs.

One of the oldest biological databases, the PDB was founded in 1971 in a very forward-thinking move, considering that only about a dozen protein structures had been solved at the time and just a few scientists required access to the data. But by the time the RCSB replaced Brookhaven National Laboratory as custodian of the database in 1998, the PDB contained approximately 8,000 entries. That number has now grown to more than 18,000. Today, the demand for structural data permeates every biological field — from antibiotic resistance in bacteria to learning and memory in humans. The database is free to all on the Web ([www.pdb.org](http://www.pdb.org)) and has become an indispensable resource for biologists of all stripes.

### PDB for Free

The Protein Data Bank

is free to search on the

Web at [www.pdb.org](http://www.pdb.org).

Biologists can deposit

data, download files,

beta-test new features,

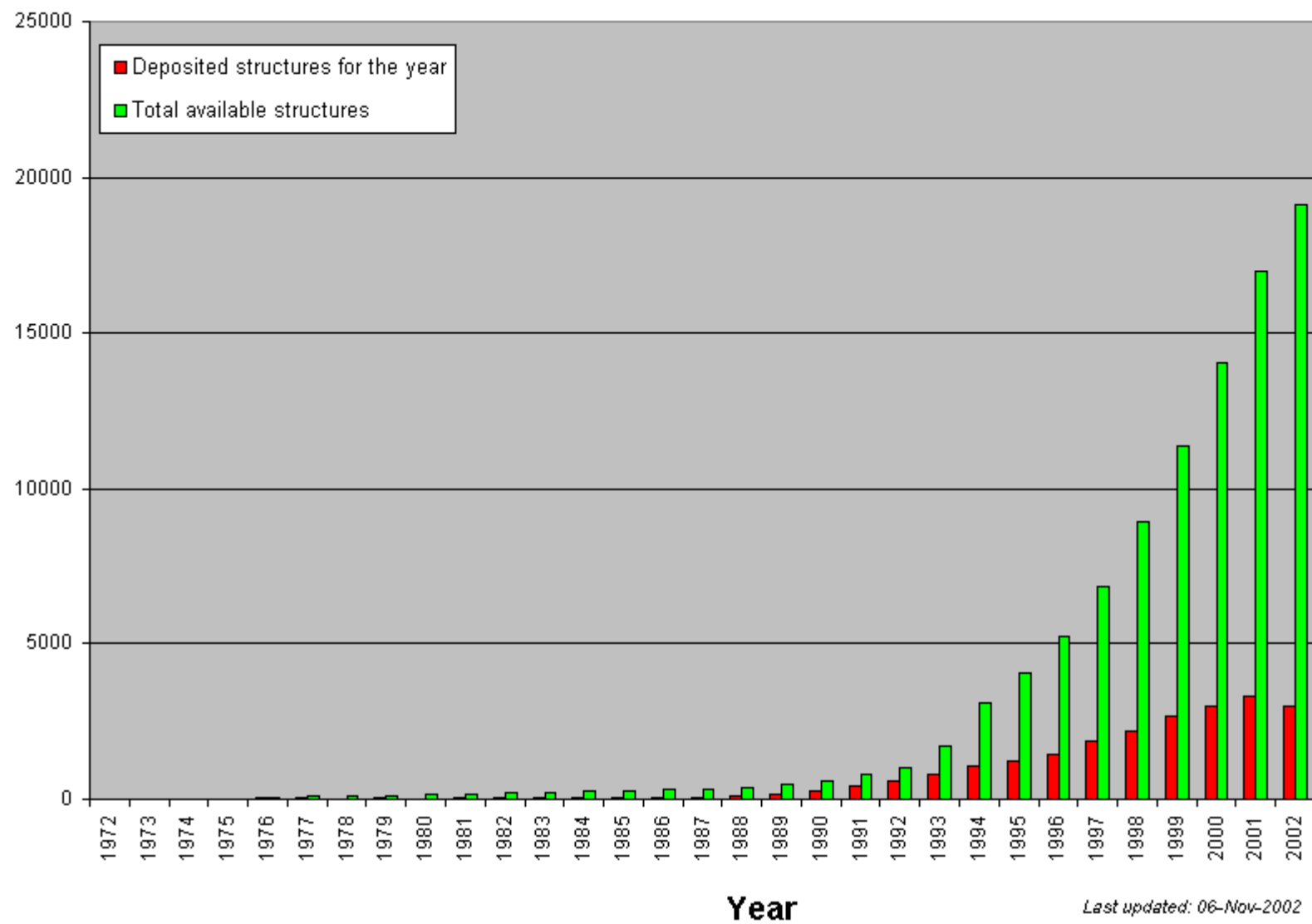
and check the status of

unreleased protein

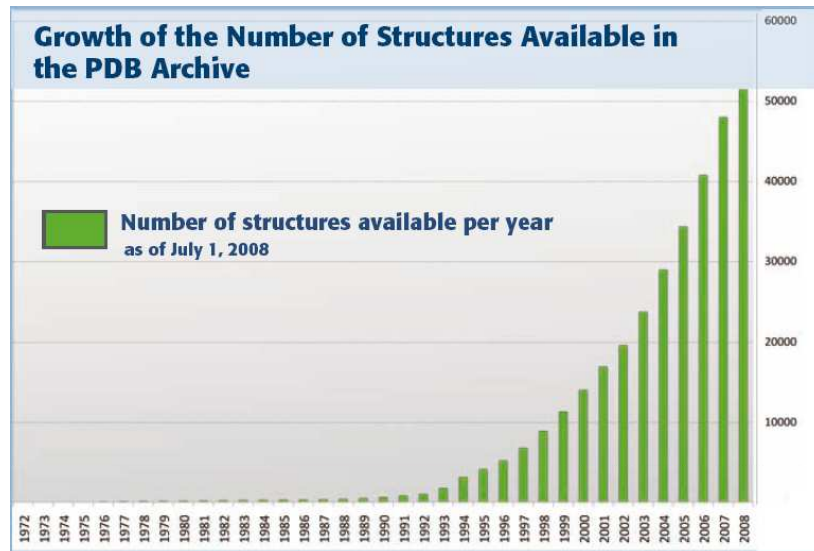
structures at the site.

Reprinted by Reprint Services with permission from BioIT World. Go to [www.bio-itworld.com](http://www.bio-itworld.com) for more information. For reorders call Reprint Services 651.582.3800. For subscription information call 508.628.4809. (Deborah Walsh)

227065



# PDB nel 2008

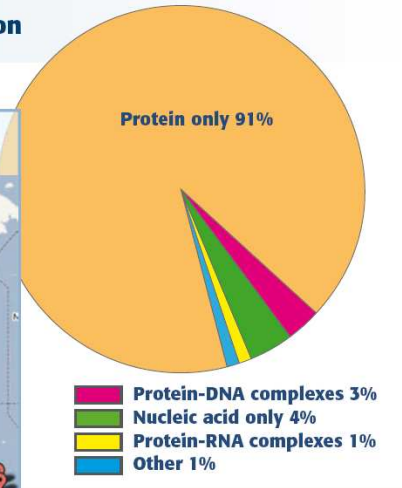


e status of



Image created using Google Earth.<sup>18</sup>

## Structure Distribution



Distribution of proteins and nucleic acids in the archive in 2008

# PDB nel 2009

## PDB Current Holdings Breakdown

		Molecule Type				
		Proteins	Nucleic Acids	Protein/NA Complexes	Other	Total
Exp. Method	X-ray	31039	933	1433	28	33433
	NMR	4952	730	122	6	5810
	Electron Microscopy	94	10	33	0	137
	Other	77	4	3	0	84

RCSB Protein Data Bank - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.rcsb.org/pdb/static.do?p=general\_information/news\_publications/index.html

Universit  My Sites Angiogenesis Ricerca Mail Database JCR-Web PDB Sequence Search Volagratis Google Maps Banca di Roma

Disable Cookies CSS Forms Images Information Miscellaneous Outline Resize Tools View Source Options

Structural O... Identifying t... CDK Descrip... Weka 3 - Da... Milano Chem... ChI&QSARS... Index of ftp... NIST NIST Scientif... http://...shtml/ QSAR World... QSAR World... Partial Least... Partial Least... Free Statisti... TEMP PLSIN... QOoConv - ... svetnik 3d q... JN CRAN at So... I-TASSER re... cdc25a in Un... 2310022A10... DOVIS - M... BHSAT docking.org... BMRB - Biolo... Design, Synt... QSAR: Predi... wade QSAR WOMBAT Protein-ligan... nendoU hom... (IUCr) Cryst... Crystal struc... ScienceDirec... Synergies of... Fast structu... Fast Structu... fred docking... PLoS Compu... VLife Techno... Effects of pr... Comparative... Divisione di... Nuove Prosp... Wiley InterS... Loading... http://...o.nsf/ Downloads...

(Untitled) Riptirinare... Page Load E... Registrazion... RCSB P...

File Modifica Visualizza Cronologia Segnalibri Strumenti 2

http://sw.16.im.med.umich.edu/databases/pdbbind/index.jsp pasqua 2008

Casa Conferenze Database Lavoro Macintosh Società Software & Co Ricerche Mercati on Line Journals GateWay

Synthesis ... Synthesis ... RCSB PDB ... tructural B... W Flavine ade... RCMD & 3... http://...html Toolboxes ... PDBBi...

**PDBbind**

The PDBbind Database  
Current version: 2007  
Total entries: 3,214

Home | Enroll | Admin | Contact Us | Log off |

Email  
demo@industry.com

Password  
\*\*\*\*

Forgot password?  
Log in

Browse

Search

Structure

Download

Deposit

PDB ID 10GS

Quick Search

### Welcome to the PDBbind Database!

The PDBbind database is developed in Dr. [Shaomeng Wang's group](#) at the [University of Michigan](#). The PDBbind database is designed to provide a collection of experimentally measured binding affinity data (*K<sub>d</sub>*, *K<sub>i</sub>*, and *IC<sub>50</sub>*) exclusively for the protein-ligand complexes available in the [Protein Data Bank \(PDB\)](#). All of the binding affinity data compiled in this database are cited from original references. [Click here for a brief introduction to the PDBbind database.](#)

**Current release.** The current release is version 2007, which means that it takes account for the protein-ligand complexes officially released by PDB before Jan 1st, 2007. In summary, this release provides binding data of 3,214 protein-ligand complexes. For each entry in this set, a table summarizing its basic information is available on our web site. A total of 1,300 protein-ligand complexes are selected to form the "refined set" with concerns on the quality of structures and binding data, which is compiled particularly for docking/scoring studies. For each entry in this set, properly processed structure files are provided so that they can be readily utilized by molecular modeling programs. Our web site also supports interactive substructure / similarity search within this data set. Another "core set" provides a non-redundant sampling of the "refined set". Currently, a total of 70 different proteins and 210 entries are included in this set.

**Availability.** Academic/governmental users may access the PDBbind database for free under a license agreement. Industrial/commercial users may access this database with a moderate amount of subscription fee. Please follow the instructions on the "enroll" page.

### References

[1] Wang, R.; Fang, X.; Lu, Y.; Yang, C.-Y.; Wang, S. "The PDBbind Database: Methodologies and updates", *J. Med. Chem.*, **2005**; 48(12); 4111-4119.

[2] Wang, R.; Fang, X.; Lu, Y.; Wang, S. "The PDBbind Database: Collection of Binding Affinities for Protein-Ligand Complexes with Known Three-Dimensional Structures", *J. Med. Chem.*, **2004**; 47(12); 2977-2980.

◆ This page has been visited for **043023** times since May 2004.

◆ This web site is maintained by Dr. Xueliang Fang and Dr. Renxiao Wang. Latest update: June, 2007.

◆ Microsoft Internet Explorer 6.0 is recommended for viewing this web site at 1024 X 768.

◆ JAVA plug-in 1.4 or higher is required for viewing chemical structures.

