Ann Natl Acad Med Sci (India), **41(3&4)**: 83-91, 2005

Computer Modeling for Drug Development

Gita Subba Rao
Professor, Department of Biophysics, All India Institute of Medical Sciences, New Delhi-110029

SUMMARY

With the advent of powerful computers and high resolution graphics, the time scale of drug discovery and development, as well as the cost of development of new drugs has been reduced very significantly.

We present here a bird's eye-view of the essential steps involved in the drug discovery and development process starting from identification and characterization of the drug target, the search for a 'hit' or a 'lead compound', the detailed study of the binding of the lead compound to the target, the strategy to improve and optimize the binding, and the final iterative cycle of refinement and testing to yield a suitable drug.

Finally, we describe briefly some of our research efforts in developing novel inhibitors of HIV-1 reverse transcriptase, and of protein tyrosine phosphatase 1B (PTP1B), a new drug target for type 2 diabetes and obesity.

With the advent of powerful computers with high resolution graphics and molecular modeling software packages, the time scale of drug development and the cost of development of new drugs has been reduced very significantly. Fig. 1 gives and idea of the time scale for the development of a drug by traditional methods. Through the use of computers, the initial process of screening, say 5000 compounds, can be reduced from 8-10 years to 8-10 weeks! Therefore, it is not surprising that computer modeling is now an essential and integral component of the pharmaceutical industry.

THE DRUG DISCOVERY PIPELINE

Let us start with the essential steps involved in the drug development process. Fig. 2 gives the flowchart of the process. The first step is the identification of the target based on the biological and biochemical data for a given medical condition. The target could be a known target, and the aim therefore would be to design new and more efficient drugs for it, or it could be a new target discovered through the genomics initiative, for which the active site if it is an enzyme, or the binding site if it is a receptor, has been characterized.

Correspondence: Prof. Gita Subba Rao, Department of Biophysics, All India Institute of Medical Sciences, New Delhi-110029, E-mail: gitarao@aiims.ac.in; delivered at the time of NAMS Pre-Conference CME at Mumbai-2005

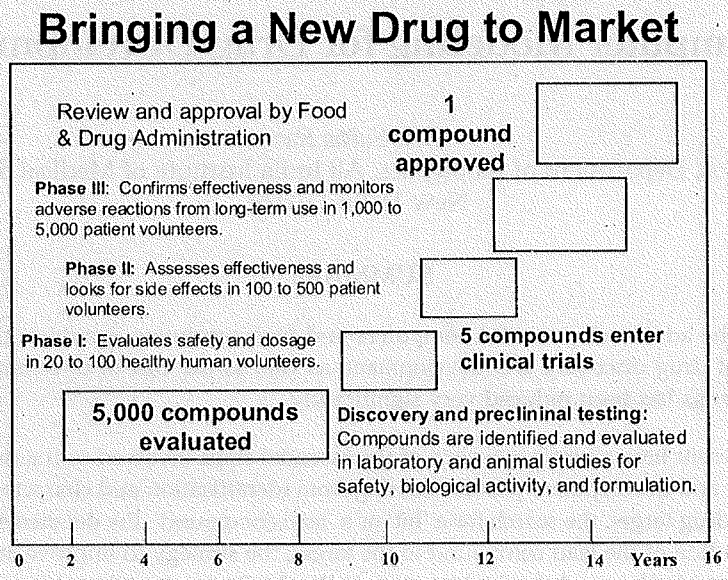

Figure 1

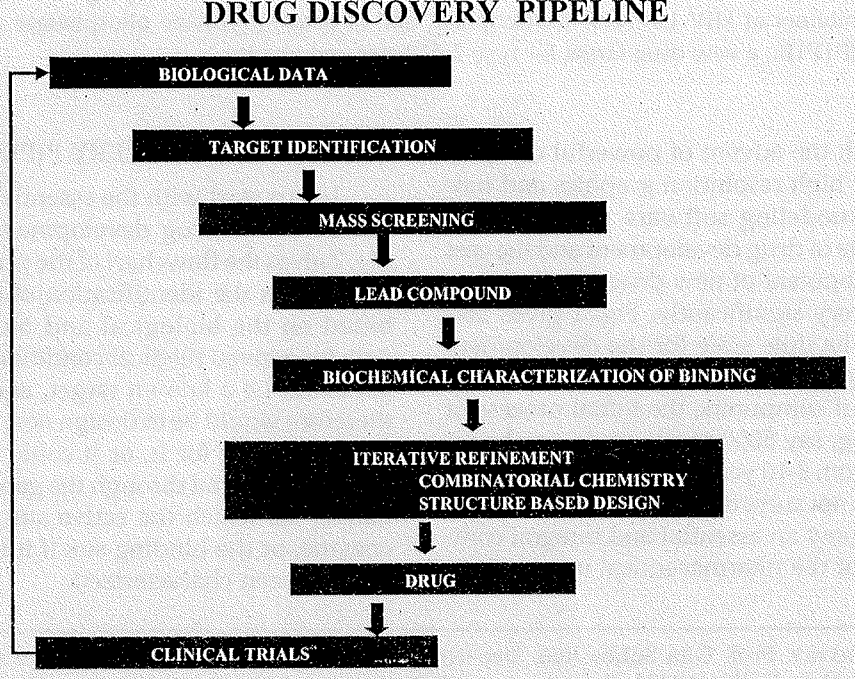

Figure 2

Once the binding region is known, the next challenge is to find suitable small molecule ligands. This is by far the most daunting task of the drug design process and can be achieved through the following stages:

1. **Mass screening**

A database of hundreds of thousand of compounds is screened by determining the activity of each of the compounds and, if a successful match is found, the initial hit is called a 'lead compound'. The lead compound is usually a weakly binding ligand and has minimal receptor activity.

In the next step of the pipeline, the binding of the lead compound to the target binding site is studied in order to determine the steric, electrostatic, hydrogen bonding and hydrophilic interactions between the ligand and the receptor. The most accurate way of doing this is through a determination of the three-dimensional structure of the complex by X-ray crystallography. A strategy can then be developed based upon the interactions to improve and optimize the binding of the lead compound. One then enters a cycle of iterative chemical refinement and testing until a drug is developed which then undergoes clinical trials. The commonly used refinement techniques are combinatorial chemistry and structure based design.

Combinatorial chemistry

This is a synthetic tool that enables chemists to rapidly generate thousands of lead compound derivatives for testing. As shown in Fig. 3, starting from a scaffold which contains a constant part and variable substituent groups, a large number of derivative structures can be generated as a result of the combinatorial process. The combinatorial libraries selected are based on the study of the binding site.

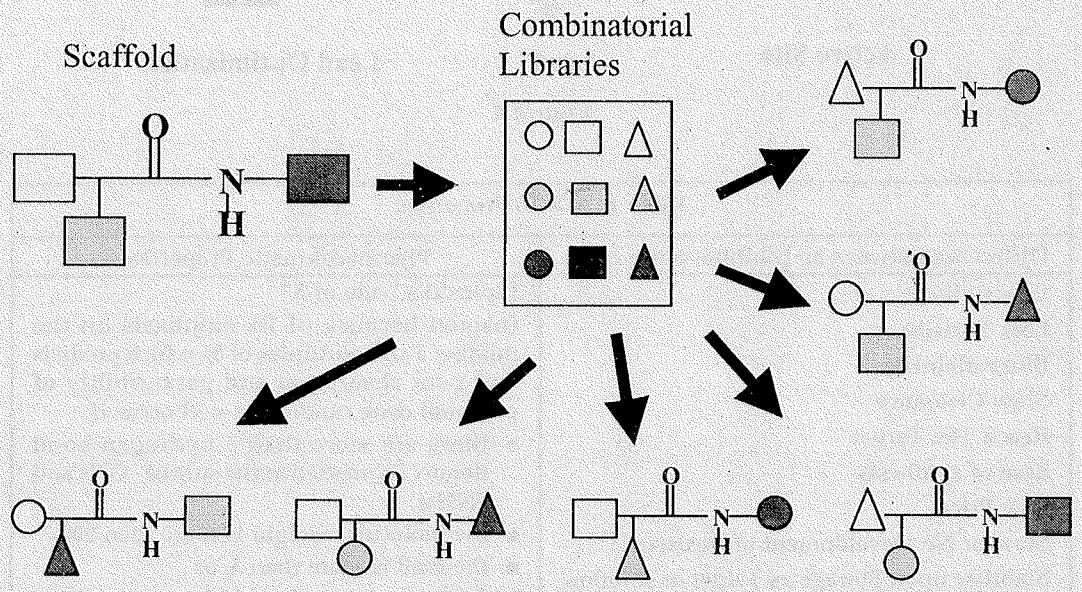

Figure 3

Structure based design

This is often called rational drug design. Based on the crystal structure of the complex, one can target regions of the ligand that fit poorly within the binding site and postulate chemical changes to improve steric and electrostatic complementarity with the receptor, as shown in Fig. 4. It is a much more focused method as compared to combinatorial chemistry.

Use of computers

The process of identifying a lead compound and refinement of the lead compound can be implemented with speed and accuracy with the help of powerful computers and molecular modeling software. Docking and scoring are the two major steps involved in computer-aided drug design. The process of docking is essentially the modeling of the complex of the receptor

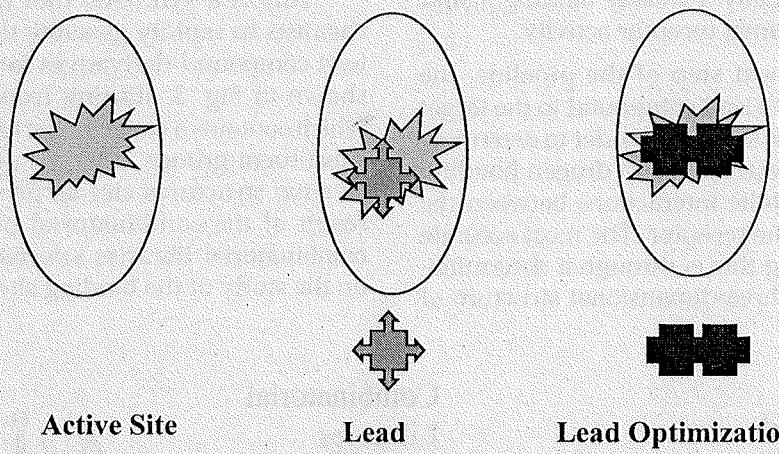

Figure 4

Table 1. ADME Properties	
Difference Between an Inhibitor and a Drug	Pharmacokinetic Properties
Selectivity Less Toxicity Bioavailability Slow Clearance Reach The Target Ease of synthesis Low Price Slow or No Development of Resistance Stability upon Storage as Tablet or solution Pharmacokinetic Parameters No Allergies	Lipinski's "rule of 5" (named because of its emphasis on the number 5 and multiples of 5), which predicts that poor absorption and permeability of potential drug candidates will occur if • there are more than 5 hydrogen-bond donors (expressed as the sum of –OHs and –NHs), • the molecular weight is more than 500, • the logP is more than 5, or • there are more than 10 hydrogen-bond acceptors (expressed as the sum of nitrogens and oxygens).

with the ligand. The latter moves around in the receptor space simultaneously changing its own structure in accordance with an energy minimization algorithm, until the optimum position is reached. Scoring is the method of predicting the affinity of binding and hence the activity of the ligand. Generally docking algorithms are combined with scoring functions to arrive at a prediction of both the ligand binding position and the activity.

Once a suitably active and selective ligand has been identified, the next step is to check whether it would be a suitable drug in terms of its ADME (absorption, distribution, metabolism and excretion) properties (Table 1). Computers play a major role in predicting these pharmacokinetic parameters, using quantitative structure-activity relationship (QSAR) methods. The basic principle of these methods is to build a model using a database of known compounds with known parameters and activities, and then to use the model to predict the parameters for an unknown compound.

Thus, given the high resolution structure of the binding site, the entire drug development process leading up to the stage of animal and clinical trials can be accomplished with the help of computers.

We now present some of our research work on the structure based design of novel inhibitors of HIV-1 reverse transcriptase and of protein tyrosine phosphatase 1B, a new drug target for type 2 diabetes mellitus and obesity.

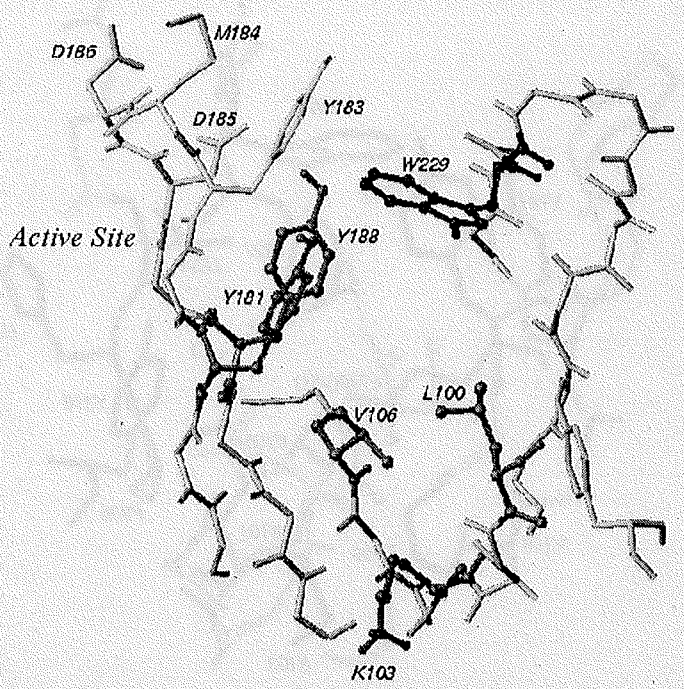

Figure 5. RT Binding pocket. Conserved residue is shown in black, frequently mutating residues in dark gray.

HIV-1 reverse transcriptase

This is a key enzyme in the life cycle of HIV-1, and is one of the targets of HAART. The main problem encountered with anti-HIV drugs is the rapid emergence of drug-resistant mutations and, therefore, the development of new, mutation resilient drugs presents a major challenge in anti-HIV therapy.

The X-ray crystal structure (1) of the complex of HIV-1 reverse transcriptase (RT) with a known inhibitor, nevaripine, indicates a binding pocket which is close to, but not at the active site of RT (Fig.5). This binding pocket has some residues that are conserved and others that mutate in response to binding of the drug. The strategy for new drug development is that the inhibitor should have a large number of interactions with the conserved residues and the backbone of the binding pocket in order to have a more mutation resilient compound. This has given rise to many "second generation" inhibitors, the best amongst them being the compound S-1143. The crystal structure of the complex with S-1143 (2) indeed shows a large number of hydrogen bonds and interactions with the backbone. A second prerequisite for an inhibitor to be potent and mutation resilient is from the analysis of the crystal structures of complexes with mutant RT's, which show that the inhibitor should be able to adapt to a mutated binding pocket.

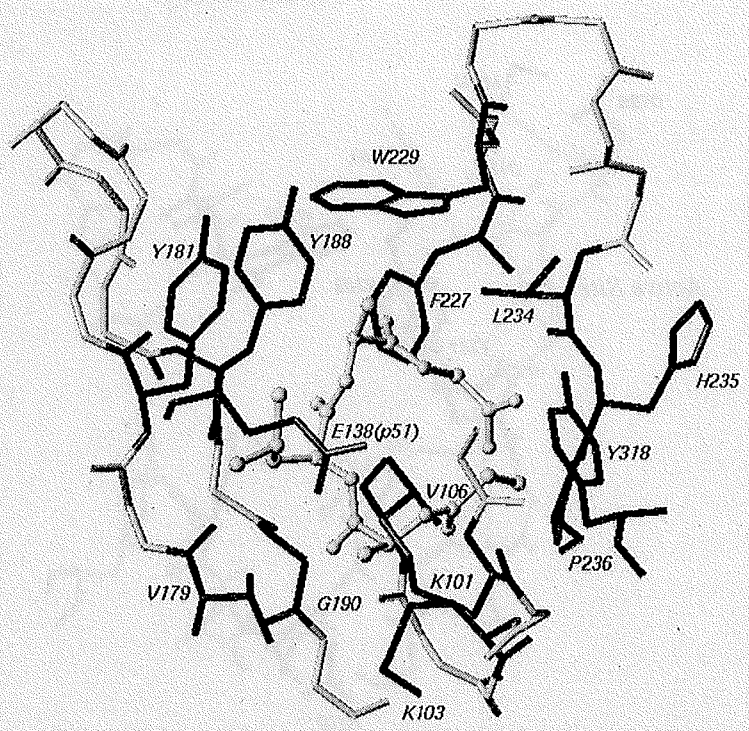

Figure 6. Final docked position of the designed peptide (in ball-and-stick rendering) in the RT binding pocket. The interacting residues are shown in black.

These observations prompted us to design a novel, small peptide inhibitor (3). The energy minimized structure of the peptide was computed, followed by docking studies, and the results are shown in Fig. 6. The peptide is found to have a large number of hydrogen bonds and interactions with the backbone of RT. The peptide also has interactions with the conserved residue W229. In addition, docking with mutant RT's shows that the hydrogen bonds and interactions are retained and, more significantly, the peptide is seen to flip over in order to adapt itself to the mutant RT pocket. Tables 2 and 3 indicate that the designed peptide inhibitor is as potent as the existing inhibitors. Thus, we have identified a new lead compound which is both potent and mutation resilient.

Table2. Measured Activities and molecular weights of known inhibitors

Inhibitor	Activity	Mol. Wt.
Nevirapine	0.08μM	266.3
UC-781	0.009μM	335.0
S-1153	0.01μM	451.4

Table 3. Predicted Activities

Ligand	Docked with	Activity (μM)
Nevirapine	Wild type RT	0.05
Designed inhibitor	Wild type RT	0.01
Designed inhibitor	Y188C mutant	0.06
Designed inhibitor	K103N mutant	0.01

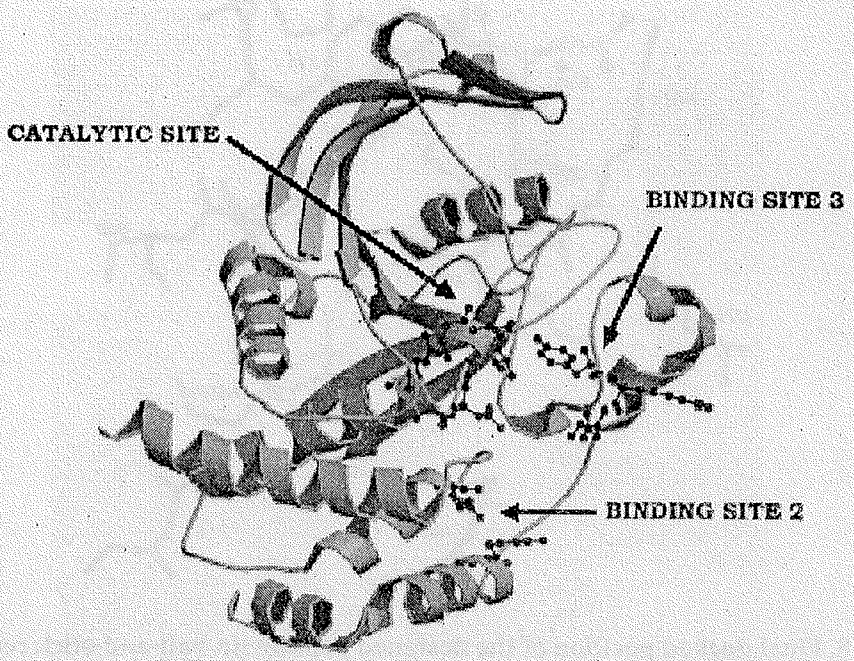

Figure 7. Crystal structure of PTP1B showing the three binding sites.

Protein Tyrosine Phosphatase 1B (PTP1B)

PTP1B has been shown to be a negative regulator of insulin signaling by dephosphorylating key tyrosine residues (4). Recent clinical studies have indicated a correlation between a polymorphism in the PTPN1 gene which encodes PTP1B, and the risk for type 2 diabetes mellitus (5). In addition, recent gene knockout studies in mice have shown the mice to have increased insulin sensitivity and improved glucose tolerance, as well as a resistance to diet-induced obesity (6). Thus there is compelling evidence that selective, small molecule inhibitors of PTP1B may be effective in treating insulin resistance at an early stage thereby leading to a preventive strategy for type 2 diabates and obesity.

The crystal structures of the complexes with known inhibitors reveal that in addition to the phosphotyrosine binding site, there are two other binding sites (Fig. 7).

Inhibitors that bind to these additional binding sites are found to have a high potency and a good selectivity for PTP1B.

Based on the above characteristic features, we have designed a novel, small peptide inhibitor. The modeled complex is shown in Fig.8.

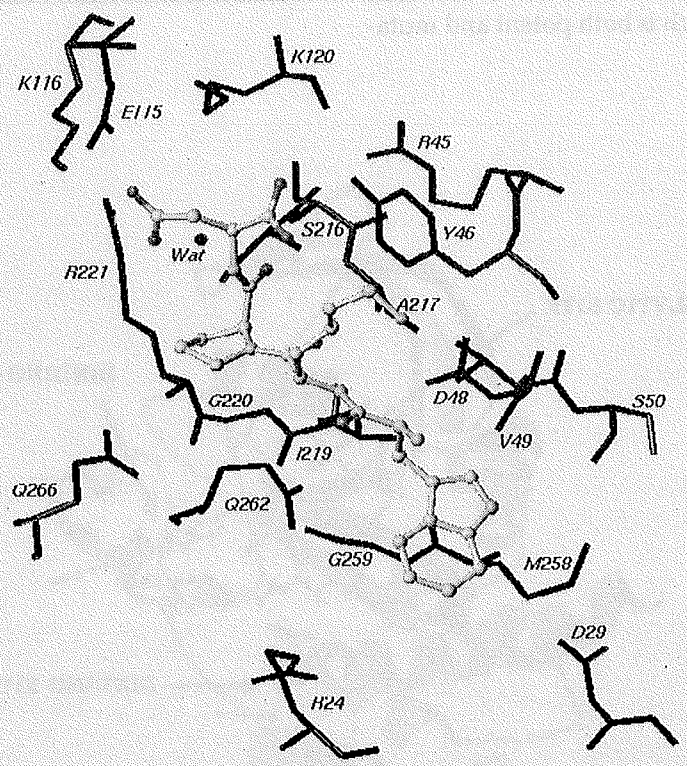

Figure 8. Final docked position of the designed peptide (in ball-and-stick rendering) in the PTP1B binding site. The interacting residues are shown in black.

The peptide is found to have hydrogen bonding with residues in the catalytic site as well as with those in binding site 3, and has interactions with residues in binding site 2. The predicted activity is 1.03 nanomolar, which is comparable to that of the most potent known inhibitor.

The next crucial step is to check whether the designed inhibitor is selective for PTP1B as compared to other closely related protein tyrosine phosphatases such as LAR, calcineurin and the highly homologous T-cell protein tyrosine phosphatase (TCPTP). This was done by modeling the respective complexes and calculating the activities. Table 4 shows that the designed inhibitor is highly selective for PTP1B. Significantly, there is a 800-fold selectivity over TCPTP. This is much higher than the best selectivity of 10-fold achieved so far by the 'best' inhibitor.

Table 4. Predicted Activities

Target	Activity (M)	Selectivity ratio
PTP1B	1.03×10^{-9}	1
TCPTP	8.34×10^{-7}	809
LAR	7.16×10^{-7}	696
Calcineurin	1.64×10^{-7}	159

Thus, the designed small peptide inhibitor is predicted to have a potency comparable to the most potent known inhibitor and a high selectivity over closely related protein tyrosine phosphatases, and is therefore, a suitable lead compound for the development of new drugs for obesity and type 2 diabetes mellitus.

REFERENCES

1. Ren J, Esnouf R, Garman E et al. (1995). High Resolution structures of HIV-1 RT from four RT-inhibitor complexes. *Nat. Struct. Biol.* **2:** 293-30.
2. Ren J, Nicholas C, Bird L et al. (2001). Structural mechanisms of drug resistance for mutations at codons 181 and 188 in HIV-1 reverse transcriptase and the improved resilience of second generation non-nucleoside inhibitors. *J Mol Biol* **312:** 795-805.
3. Subba Rao G and Bhatnagar S (2003). *In Silico* structure-based design of a potent, mutation resilient, small peptide inhibitor of HIV-1 Reverse Transcriptase. *J. Biomol Struct. Dynamics* **21:** 171-178.
4. Ukkola O and Santaniemi M (2002). Protein tyrosine phosphatase 1B: A new target for the treatment of obesity associated co-morbidities. *J. Int. Med* **251:** 465-475.
5. Palmer ND, Bento JL, Mychaleckyj JC et al. (2004). Association of protein tyrosine phosphatase 1B gene polymorphisms with measures of glucose homeostasis in Hispanic Americans. *Diabetes* **53:** 3013-3019
6. Elchebly M, Payette P, Michaliszyn E et al. (1999). Increased insulin sensitivity and obesity resistance in mice lacking the PTP1B gene. *Science* **283:** 1544–1548.