

Diabetes & Drug Design

**Structural & computational biology
applied to a medically relevant problem**

Dr. K.A. Watson
University of Oxford
Laboratory of Molecular Biophysics

Research Themes

Study of intra- and extra-cellular signalling processes that lead to insulin production and blood glucose lowering

To understand protein-ligand and protein-protein interactions at the atomic level

Develop computational approaches that exploit known structural and biological data

Why Diabetes?

- Type II: non-insulin dependent, accounts for 90% of all diabetes
 - 10% of health care budget in UK
 - fourth leading cause of death
- Treatment
 - 50% treated with diet and drugs
- Type II diabetics show significant increased risk of stroke & heart failure
 - in addition to diabetic-related symptoms such as blindness, renal failure, & amputation

Diabetes & Drug Design

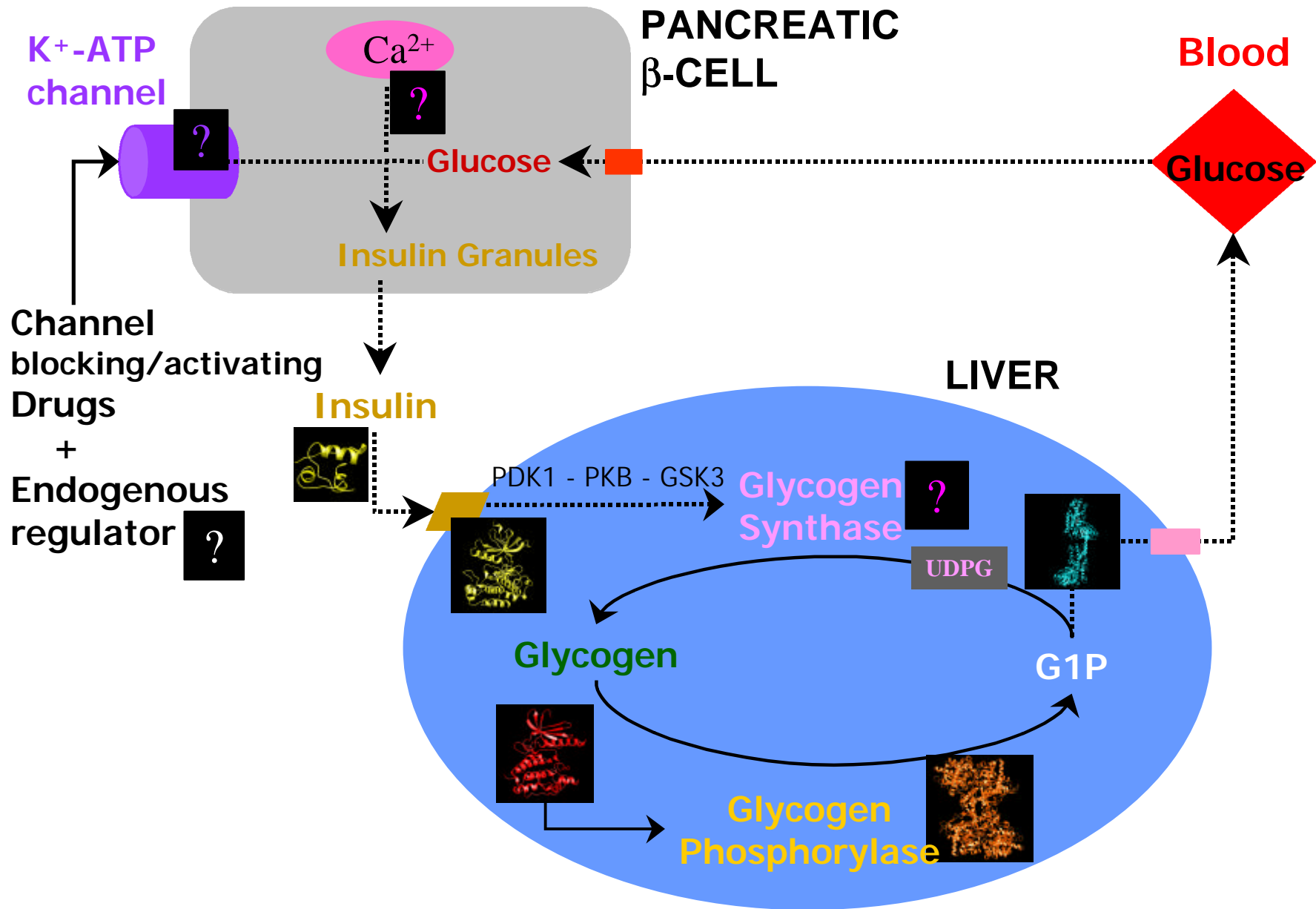
Structural Studies

- *Glycogen metabolism*
 - Glycogen phosphorylase
 - Glycogen synthase
- *Insulin secretion*
 - K-ATP channel (SUR1/Kir6.2)
 - α -endosulfine
 - Ca⁺⁺ signalling proteins

Computational Studies

- QSAR
 - GRID/GOLPE
- Free Energy Perturbation
 - TIA
- Machine Learning
 - ILP

Structural Biology



Computational Biology – WHY?

- **Useful in the absence of experimental data**
 - Validate by experiment, where possible
- **Machines are faster, less subjective**
 - better for handling large amounts of data
 - most effective method of pattern recognition
- **To achieve clearer insight into protein-ligand interaction energies**
 - Free energy perturbation (TIA)
- **To extract rules governing the affinity of a given interaction and learn the structural aspects which may be correlated to the pharmacokinetics**
 - Machine learning (ILP)

Computational Biology – HOW?

- **Development & Application of Novel Methods**

Using databases of known protein-ligand structures

- QSAR (GRID/GOLPE)

- GP inhibitor complexes

- Correlation of 3-D chemical structure with biological activity

- Thermodynamic Integration Analysis (QM/MM, $\Delta\Delta G$)

- GP inhibitor complexes (subset of 5 structures)

- Deconvolution of binding energetics

- Machine Learning (Inductive Logic Programming)

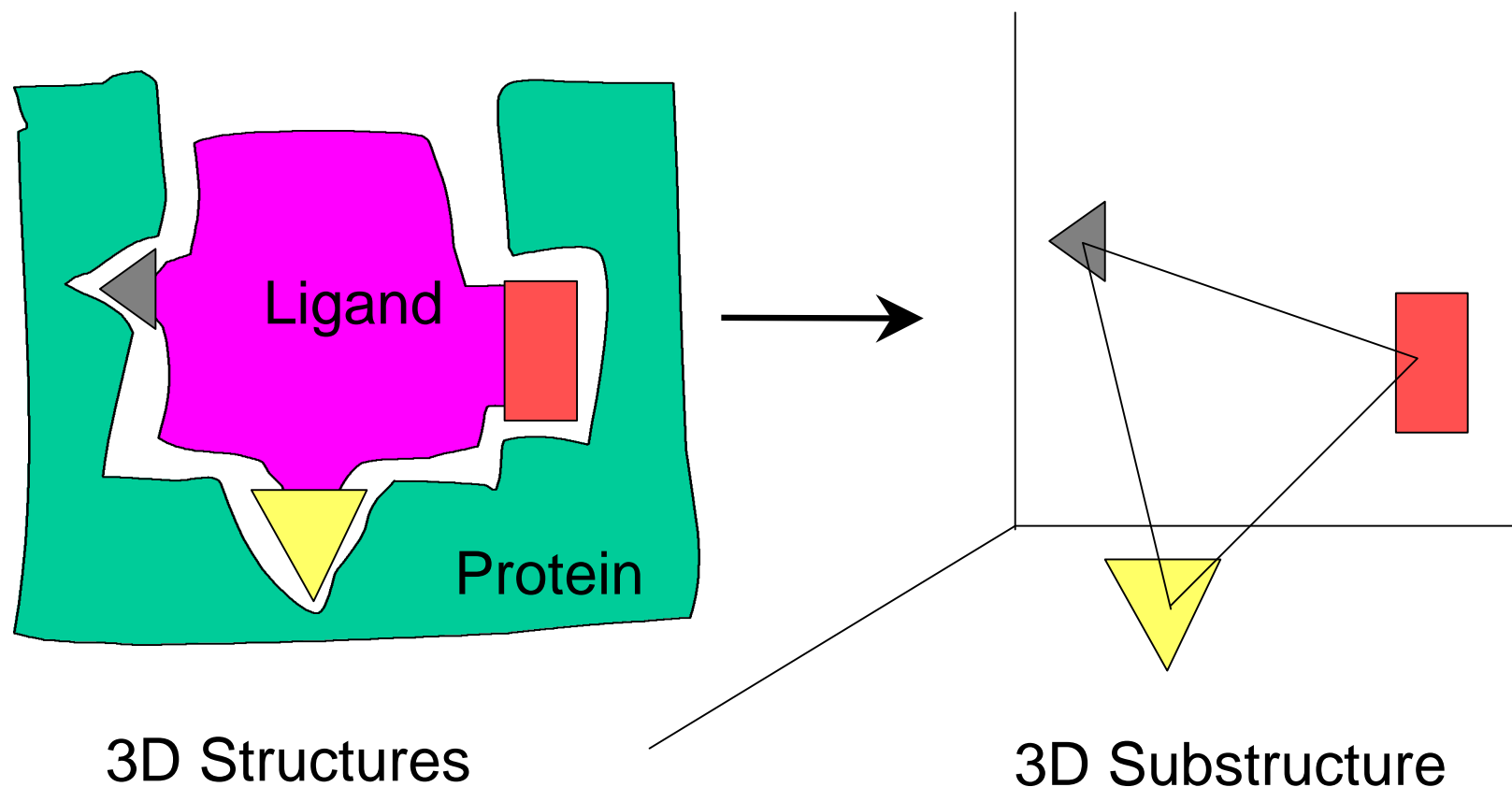
- K⁺ATP channel blockers & activators

- Pharmacophore design & QSAR

Machine Learning

**Pharmacophore design
using
Inductive Logic Programming**

Pharmacophore Design



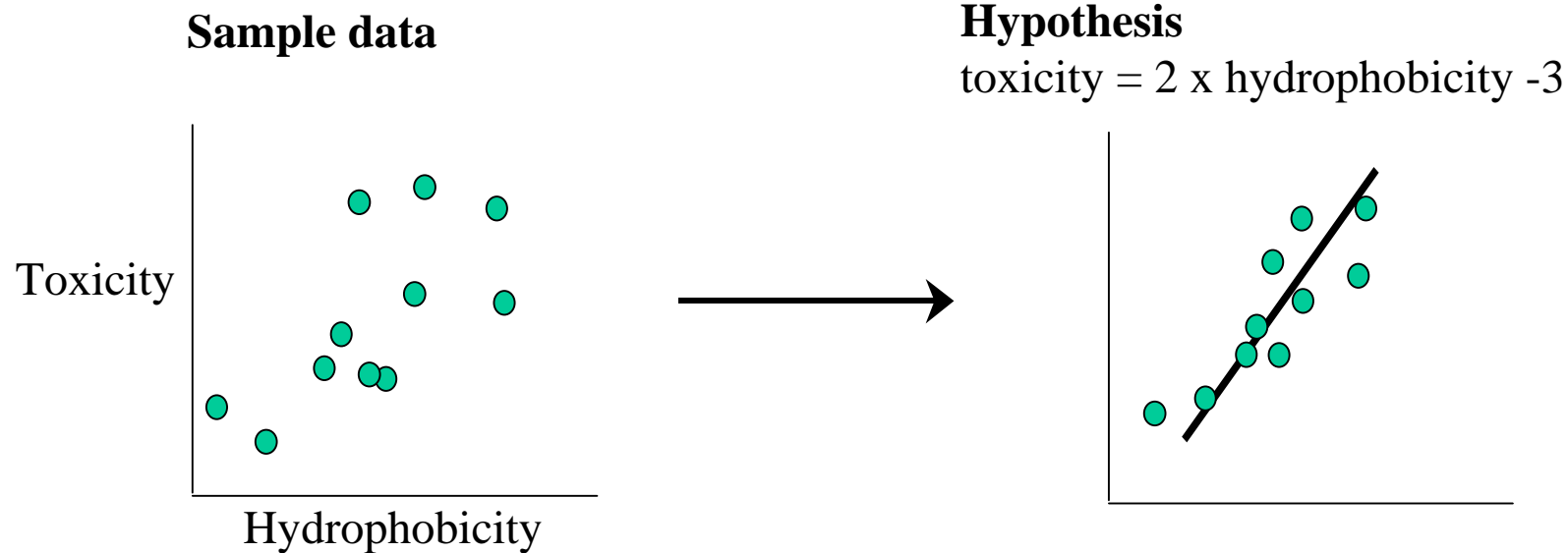
3D Structures

3D Substructure

Subset of atoms or functional groups and its 3D geometry that is responsible for a given biological activity

Machine Learning

- Programs that hypothesize general descriptions from sample data



- Inductive Logic Programming (ILP)

Deductive vs. Inductive Logic

Hypothesis

X is the grandfather of Y if X is the father of some Z,
who is a parent of Y

Background Knowledge

Henry is the father of Jane
Henry is the father of Joe
Jane is a parent of John
Joe is a parent of Robert

deductive



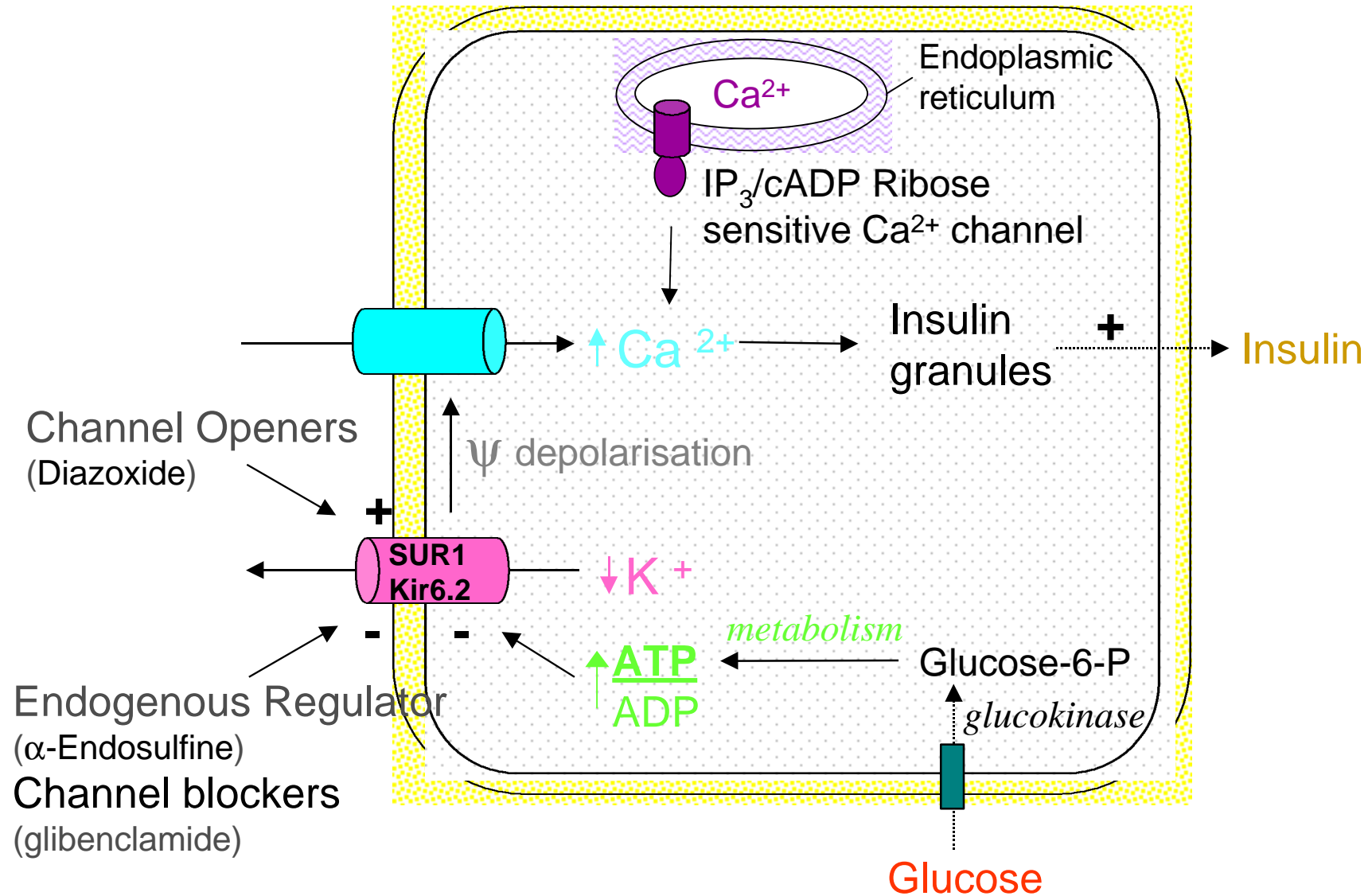
Fact/Examples

Henry is the grandfather of John
Henry is the grandfather of Robert

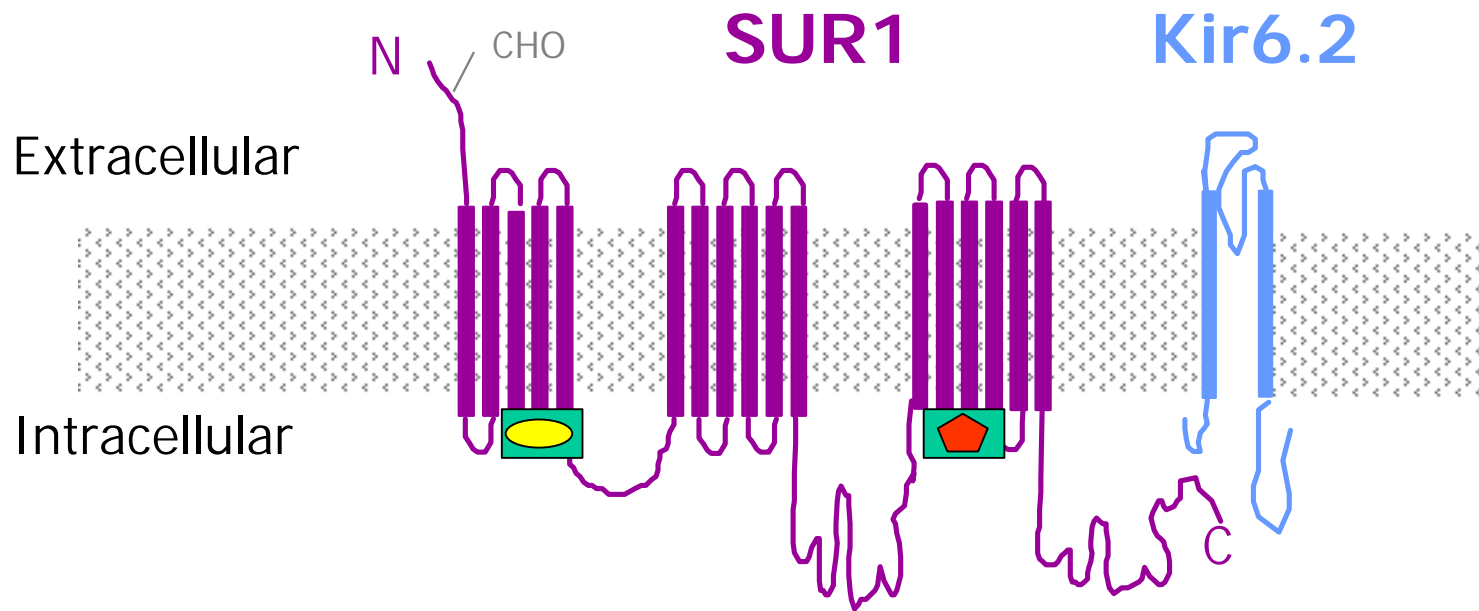
inductive






The Target: K^+ -ATP channel in the pancreatic β -cell

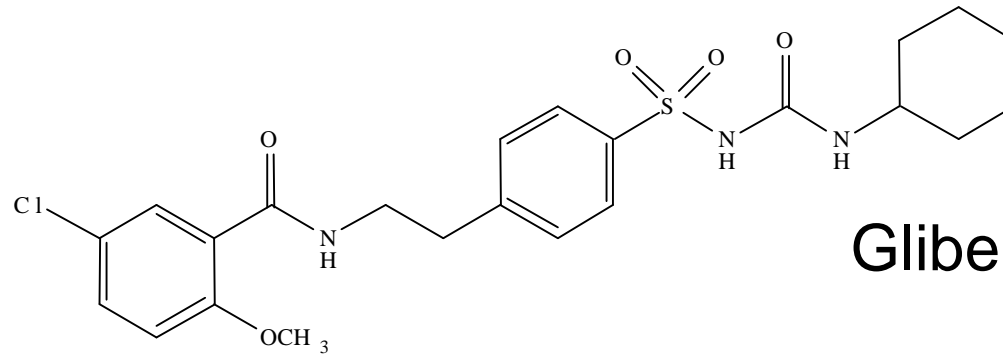


K⁺-ATP channel – the binding sites

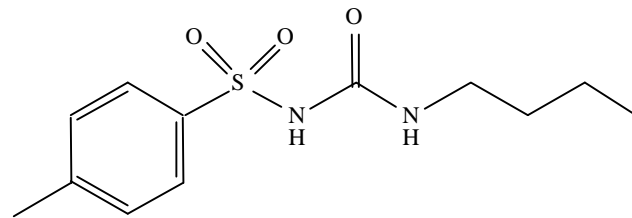


-  Glibenclamide
-  Meglitinide
-  Tolbutamide

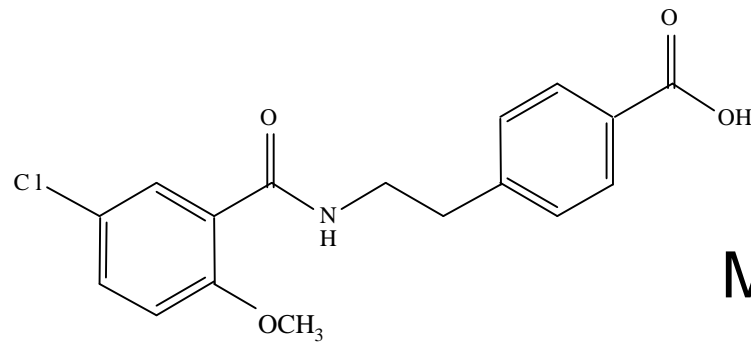
K⁺-ATP channel blockers



Glibenclamide



Tolbutamide



Meglitinide

K⁺-ATP channel and ILP

Facts/Examples:

100's

Ensemble of 3D structures of channel blocking & channel activating molecules (CSD & modelled)

+

Background:

10,000's

Pharmacophores consist of 3-5 “pieces” (geometric parameters)

General chemical knowledge (ex. H-bond donors/acceptors)



Hypothesis:

10's

A molecule is a channel blocker **if...**

A molecule is a channel activator **if...**

Representation

- Description of channel blockers

atm (m1, a1, o, 3, 5.915, -2.441, 1.799)

atm (m1, a2, c, 3, 0.574, -2.773, 0.337)

atm (m1, a3, s, 3, 0.408, -3.511, -1.314)

bond (m1, a1, a2, 1)

bond (m1, a2, a3, 1)

- Predicate for 'channel blocker'
- Predicates for hydrogen acceptors/donors, hydrophobic groups, halogens, euclidean 3D distances, etc.

Active(x):- hacc(X,A), hdonor(X,B), dist(X, A,B, 3.1, 1.1)

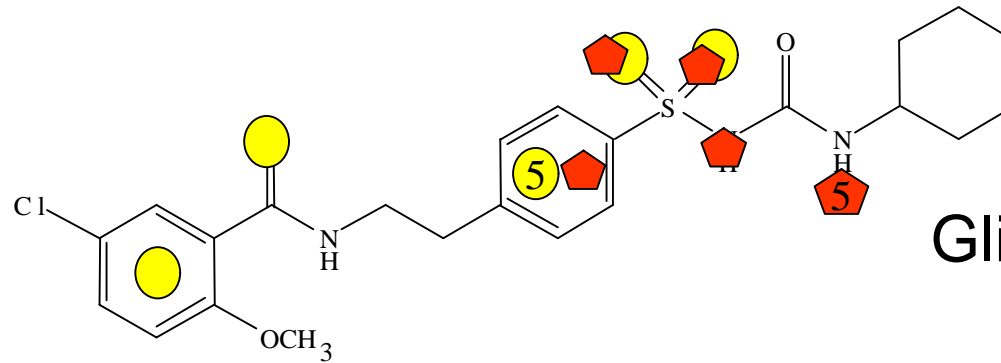
Background Knowledge

- Hydrogen donors/acceptors and hydrophobic groups are always potentially relevant
- Pharmacophores typically consist of 3-5 pieces
- The geometry of a pharmacophore is specified by the distances among the pieces, with a 0.75 – 1.0 Å tolerance for each distance
- Look for 2 pharmacophores, since there may be two distinct binding sites

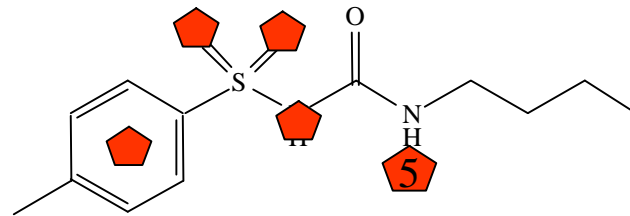
Result (Expt 1)

- **Find a 4 point pharmacophore (1.1Å tolerance) that describes all the compounds**
- **Molecule A has the desired activity in some conformation if:**
 - Molecule A contains a hydrogen acceptor C and
 - Molecule A contains a hydrogen acceptor D and the distance between C and D is $4.27 \pm 1.1 \text{ \AA}$ and
 - Molecule A contains a hydrogen donor E and the distance between C and E is $2.27 \pm 1.1 \text{ \AA}$ and
 - Molecule A contains a hydrophobic group F and the distance between D and F is $3.79 \pm 1.1 \text{ \AA}$ and the distance between C and F is $4.17 \pm 1.1 \text{ \AA}$ and the distance between E and D is $2.39 \pm 1.1 \text{ \AA}$ and the distance between E and F is $3.90 \pm 1.1 \text{ \AA}$.

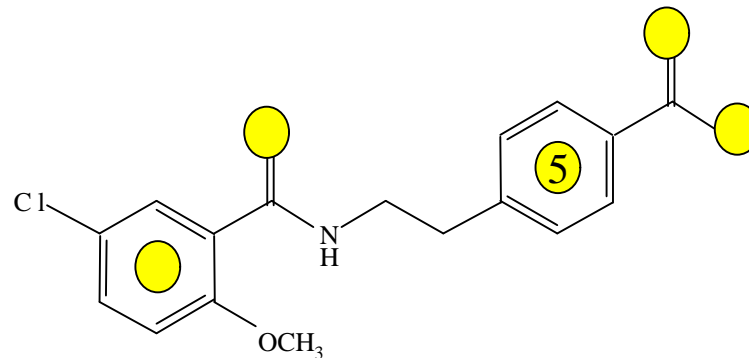
4/5 point Pharmacophores found using ILP in under a minute



Glibenclamide



Tolbutamide



Meglitinide

Utilising the Pharmacophore

- Alignment of molecules for QSAR
- Design of novel channel blockers

Work in Progress

- Quantifying the ILP approach
 - Addition of pharmacokinetic data
 - GRID/GOLPE
- K⁺-ATP channel activators

Colleagues

SUR1

Sujit Dutta
Michail Mikailov
Steve Ashcroft

α -endosulfine

Dominique Bataille (France)

CD38

Antony Galione

***E.coli* GP:HPr**

Alan Peterkofsky (Maryland)
Sujit Dutta

***C. callunae* SP**

Bernd Nidetsky (Austria)
David Hwang

GS

Reinhard Schinzel (Germany)

ILP

David Hwang
Ashwin Srinivasan
Ross King (Wales)

TIA

Georgios Archontis (Cyprus)
Martin Karplus (France)

BDA
MRC
BBSRC
Lister Institute