



# The Computational Investigations of Protein-Inhibitor Complexes: Ligand Binding Induced Fit

Francesca Toschi,<sup>1</sup>J. W. Essex,<sup>1</sup> A. R. Leach,<sup>2</sup> P. A. Bamborough<sup>2</sup>

<sup>1</sup>*Chemistry Department, University of Southampton, Southampton, SO17 1BJ*

<sup>2</sup>*GlaxoSmithKline, Gunnels Wood Rd, Stevenage, SG1 2NY*

---

# Overall aim

---

Analysis of X-ray crystal structures of protein/ligand complexes in order to:

- Characterize **side-chain conformational changes** induced by ligands in proteins;
- Identify general patterns and features of induced fit in **different protein systems**.

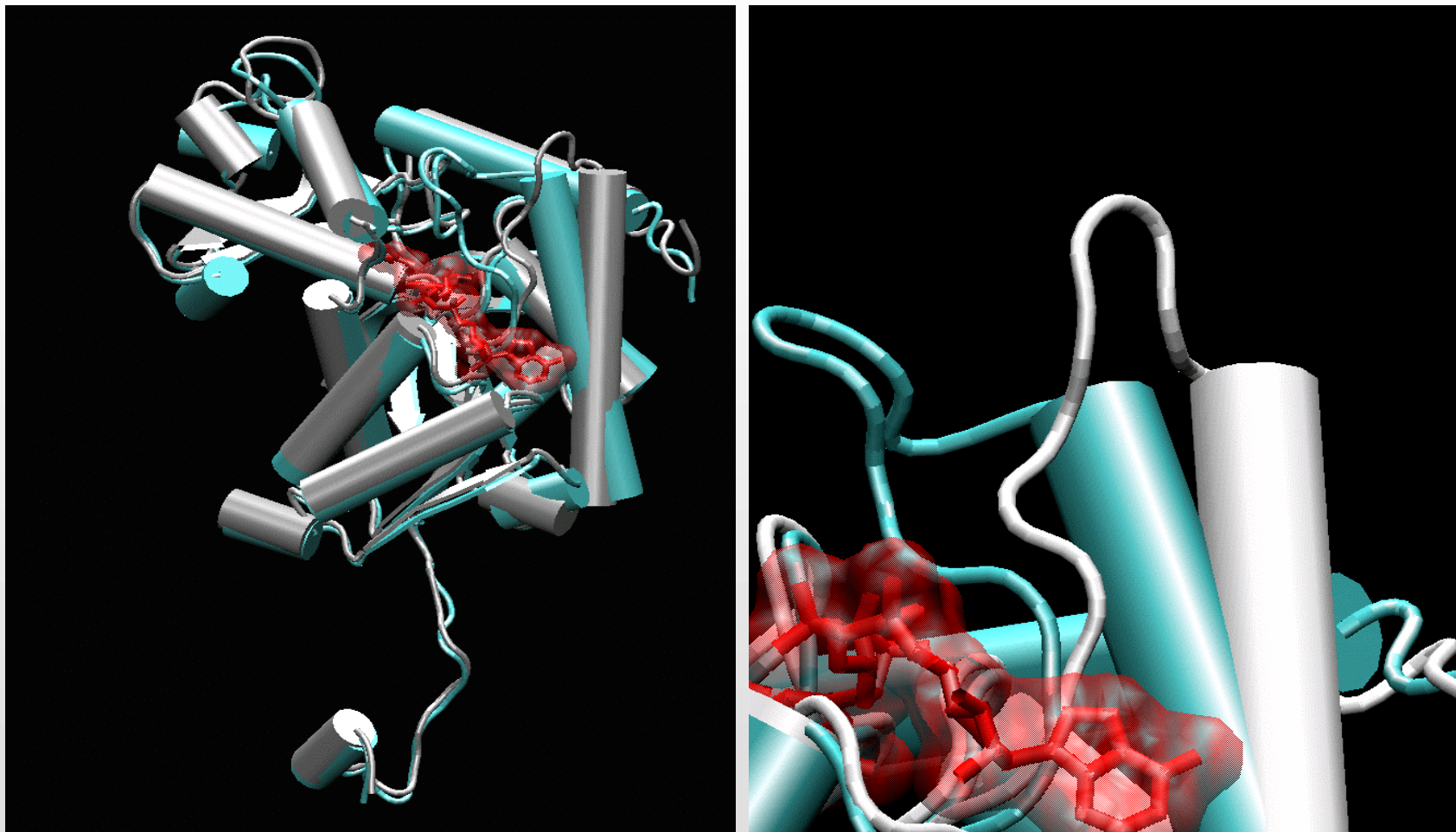
This could help the prediction of flexibilities and conformations of residues in protein binding sites.

# Background: Protein motions

**Protein motions** can be classified on the basis of their **size** and include:

- Motions of **subunits**
- Motions of **domains**
- Motions of smaller **fragments**
- Motions of **side-chains**

# Protein motions: Example



**Motion of Lactate Dehydrogenase (LDH) upon pyruvate and NADH binding.**

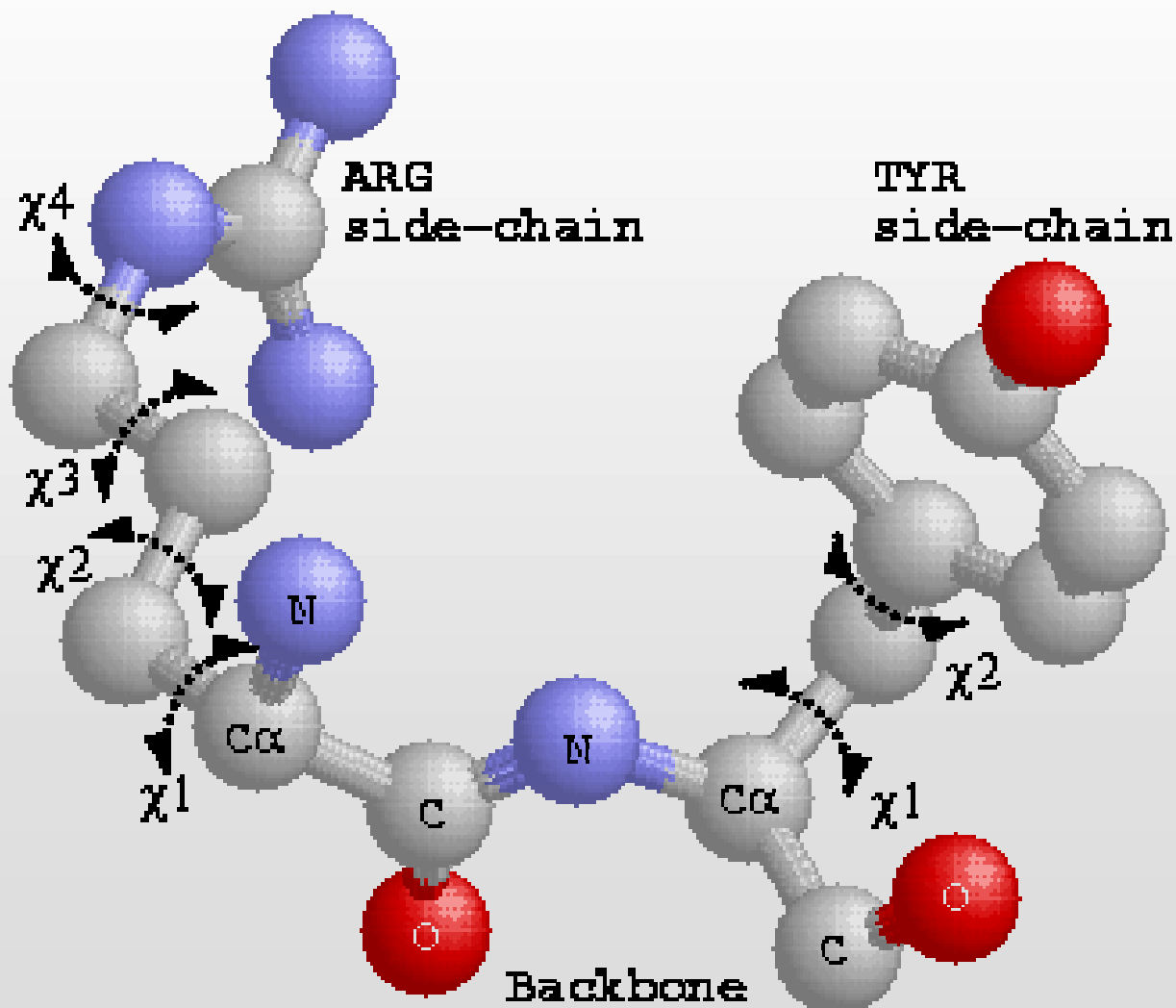
# Protein motions

## Side-chain flexibility:

- **Generally observed**
- **Prediction potentially useful for**
  - flexible docking
  - rational drug design
- **May depend on**
  - induced fit
  - intrinsic amino acid flexibility
  - crystallisation conditions and/or structure solution

# Protein motions

## Side-chain torsions



# Side-chain flexibility studies

- **Najmanovich *et al.*:**
  - **Apo-** and **holo- protein pairs** comparisons;
  - **Threshold** of **60°** in at least one torsion angle.
  
- **Zhao *et al.*:**
  - Comparisons of pairs of **uncomplexed structures**;
  - **Residue- and environment- specific confidence levels** as the angular thresholds that contain 90% of the observed  $\chi_1$  and  $\chi_2$  structural changes.

- **Flexibility scale**

- **Najmanovich**:<sup>1</sup>

**Lys > Arg, Gln, Met > Glu, Ile, Leu > Asn, Thr, Val, Tyr,  
Ser, His, Asp > Cys, Trp, Phe**

- **Zhao** (all environments):<sup>2</sup>

**Ser > Lys, Glu, Gln, Arg, Met > Leu, Asn, Asp, Val, Thr >  
Ile, His, Cys, Trp, Tyr, Phe**

[1] R. Najmanovich, J. Kuttner, V. Sobolev, and M. Eldman, *Proteins: Structure, Function, and Genetics*, **39**, 261 (2000)

[2] S. Zhao, D. S. Goodsell, and A. J. Olson, *Proteins: Structure, Function, and Genetics*, **43**, 271 (2001)

# Side-chain flexibility studies

- **Rotamer libraries:** discrete side-chain torsion angles and associated probabilities.
- **Backbone-Dependent Rotamer Library:**<sup>1</sup>
  - Side-chain angles for each rotamer type for each **10° by 10°** region of the  $\phi$  and  $\psi$  conformational space of the **backbone**;
  - Recently completed by the **Conditional Backbone-Independent Rotamer Library**<sup>2</sup> and **Bayesian statistical analysis**.<sup>2</sup>

[1] R. L. Dunbrack, and M. Karplus, *J. Mol. Biol.*, **230**, 543 (1993)

[2] R. L. Dunbrack, and F. E. Cohen, *Protein Science*, **6**, 1661 (1997)

# Side-chain flexibility studies

## Conditional Backbone-Independent and Backbone-Dependent Rotamer Libraries:

**r1** =  $\chi_1$  rotamer

**r2** =  $\chi_2$  rotamer

**r12** =  $\chi_1, \chi_2$  rotamer

for all residues having three possible  $\chi_1$  and  $\chi_2$  rotameric states, **r12 can assume 9 different states** (9 different possible combinations of r1 and r2)

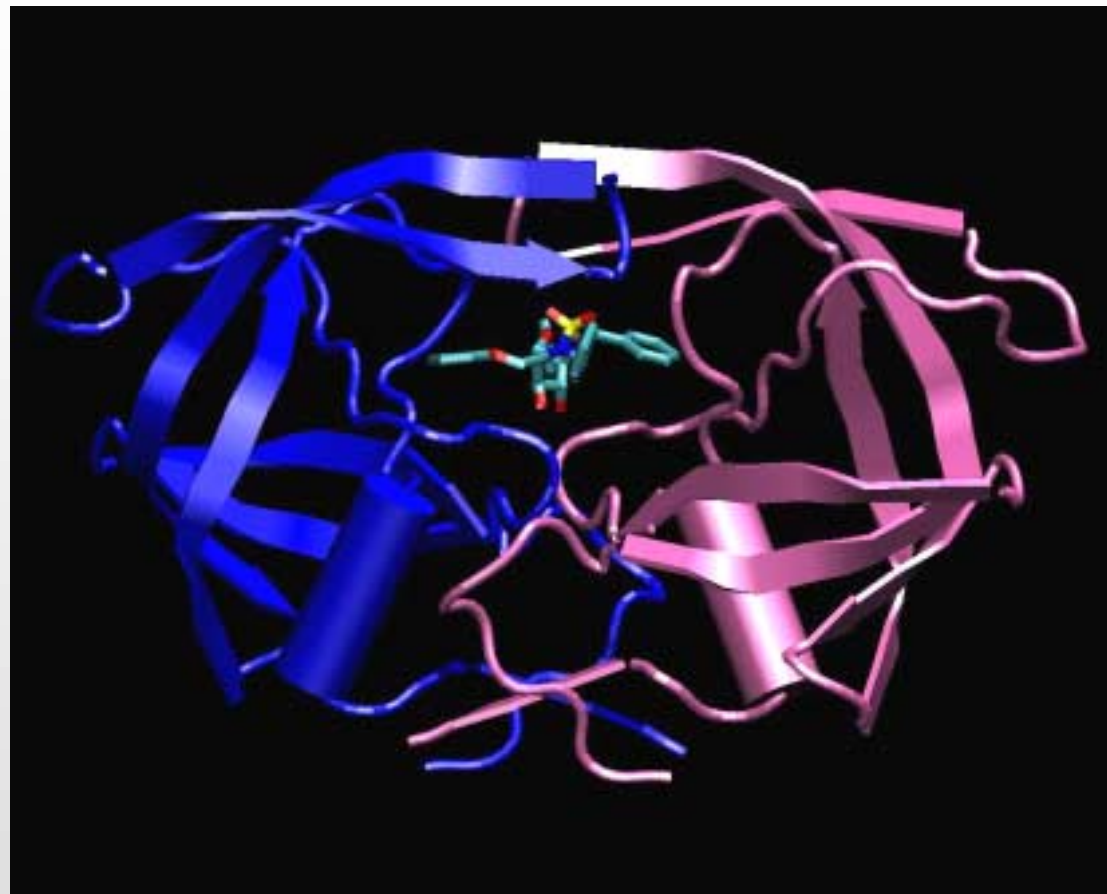
**rank** = an **integer** number ranging from 1 to 9 which describes **how common** the actual **r12** rotamer is.

# My project: aim

Looking for **systematic effects** of **ligand binding** on the **conformations** adopted by residue **side-chains**;

- Are there **conformational**, **flexibility** and/or **probability changes** that can be attributed to the effect of **ligands**?
- Can we identify **general patterns** and features of induced fit in **different protein systems**?

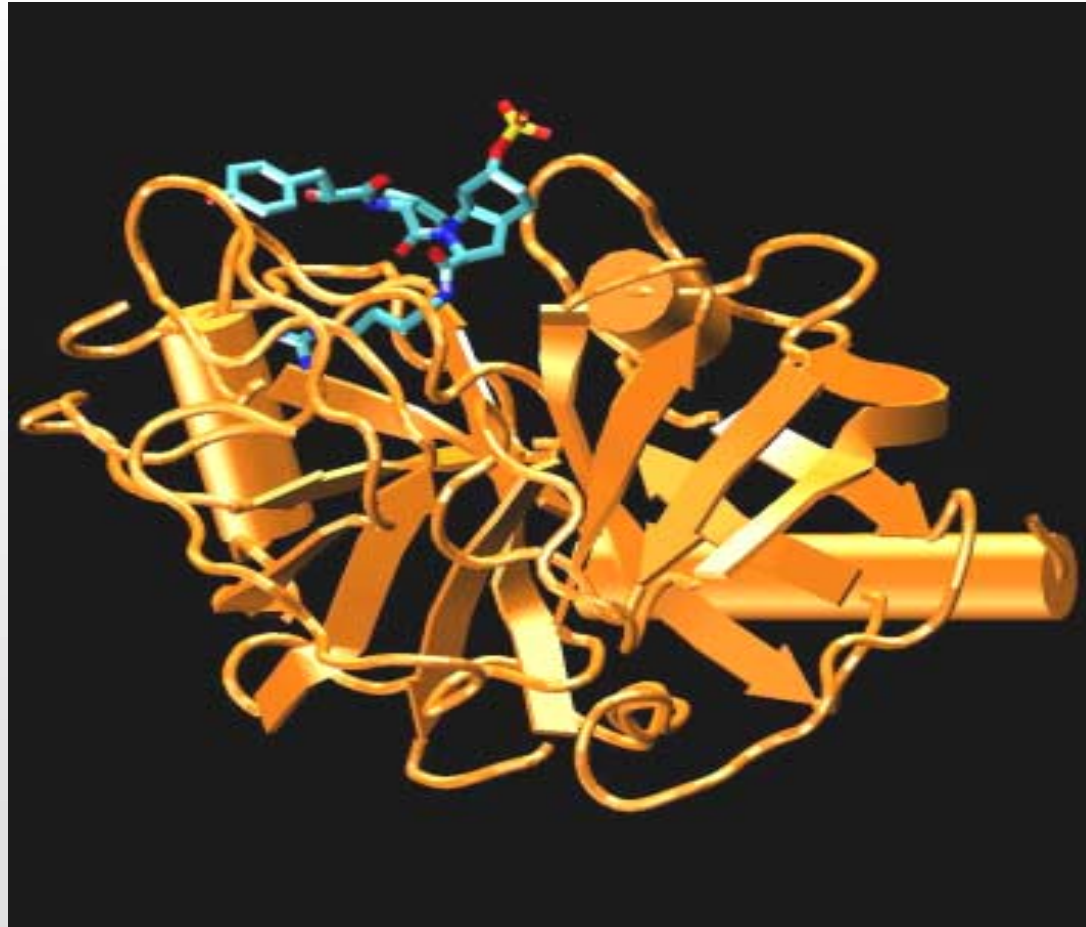
➔ In order to answer these questions, I looked at which residues move in the following dataset.



**HIV-1 protease homodimer bound  
structure (1AJV PDB file)**

## HIV-1 Protease

- 1. 24 holo-proteins**
  - resolution  $\leq 2.5 \text{ \AA}$
  - identical sequence
  - identical binding sites
  - different ligands
- 2. 17 holo-proteins**
  - subset of (1)
  - resolution  $\leq 2.0 \text{ \AA}$
- 3. Apo-protein**
  - 1G6L pdb entry (1.9  $\text{\AA}$ )



**Inhibited Trypsin (1AQ7 PDB file)**

## Trypsin

- 1. 59 holo-proteins**
  - resolution  $\leq 2.5 \text{ \AA}$
  - identical sequence
  - similar binding sites
  - different ligands
- 2. 51 holo-proteins**
  - subset of (1)
  - resolution  $\leq 2.0 \text{ \AA}$
- 3. 4 apo-proteins**
  - identical sequence
  - resolution  $\leq 2.0 \text{ \AA}$



**Inhibited Thrombin (1AE8 PDB file)**

## Thrombin

- 1. 35 holo-proteins**
  - resolution  $\leq 2.5 \text{ \AA}$
  - 3 slightly different sequences (99.7% and 99.6% sequence identity)
  - similar binding sites
  - different ligands
- 2. 13 holo-proteins**
  - subset of (1)
  - resolution  $\leq 2.0 \text{ \AA}$
- 3. 5 apo-proteins**
  - one of them having resolution  $< 2.0 \text{ \AA}$

# Methods

- Residues' torsion angles analysed employing **Najmanovich** and **Zhao** thresholds
- **Rotamers** and **probabilities** defined using the **Backbone-Dependent** and **Conditional Backbone-Independent Rotamer Libraries**
- Comparisons performed for all **apo-/apo-**, **holo-/holo-** and **apo-/holo-** protein pairs in the dataset, and for both **all residues** and only **binding site** residues.

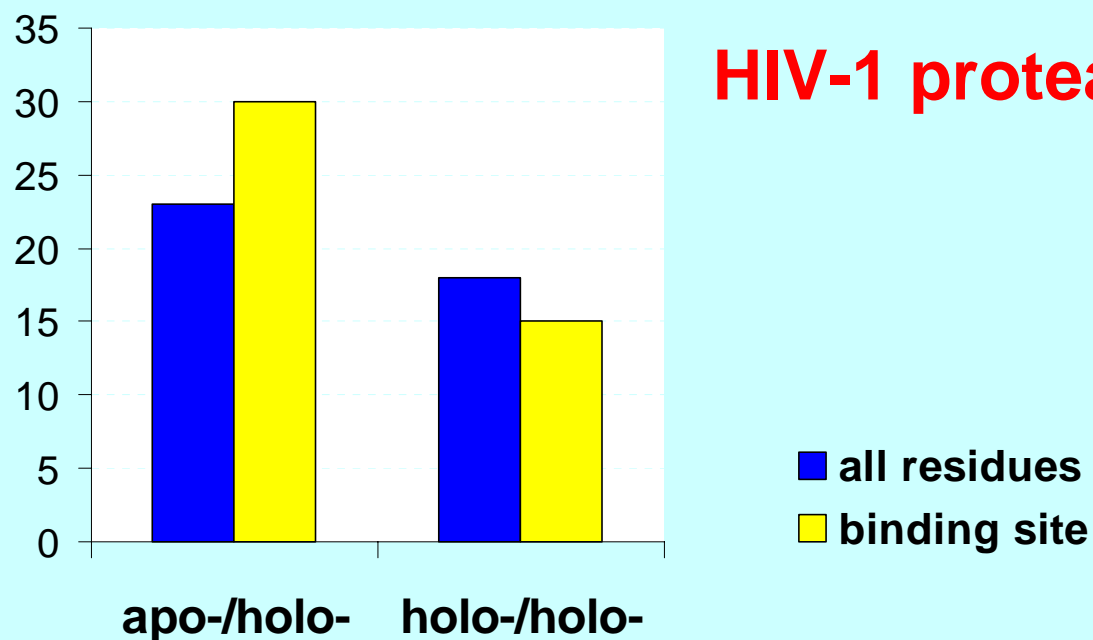
**Binding sites** defined applying the automated analysis of interatomic contacts by **Sobolev *et al.*<sup>1</sup>**.

[1] V. Sobolev, A. Sorokine, J. Prilusky, E. E. Abola, and M. Edelman, *Bioinformatic*, **15** (4), 389 (2000)

# (Some) Results

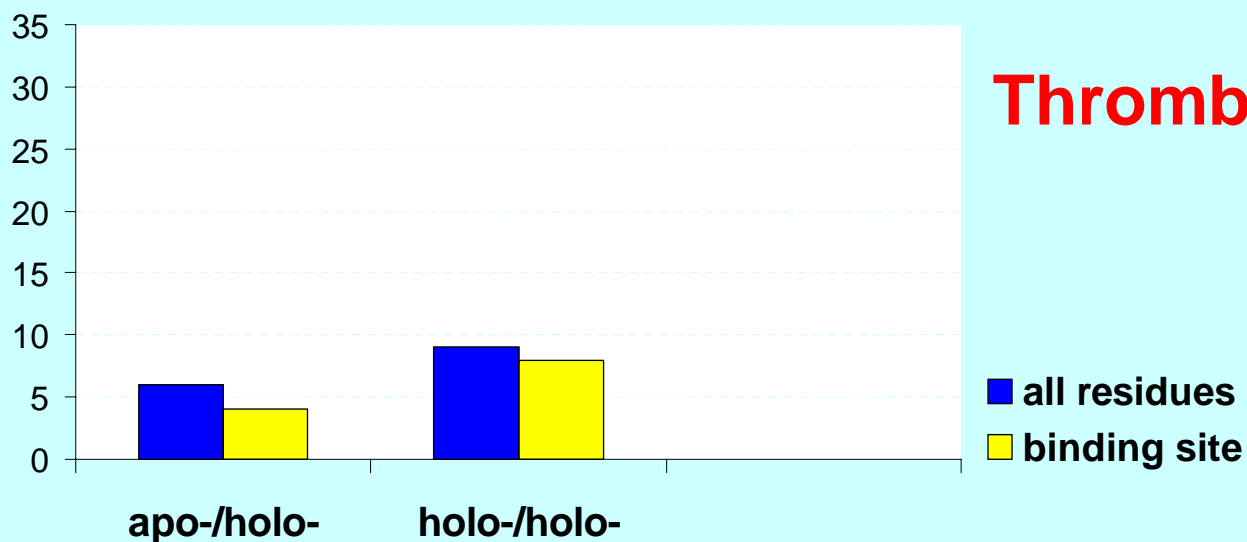
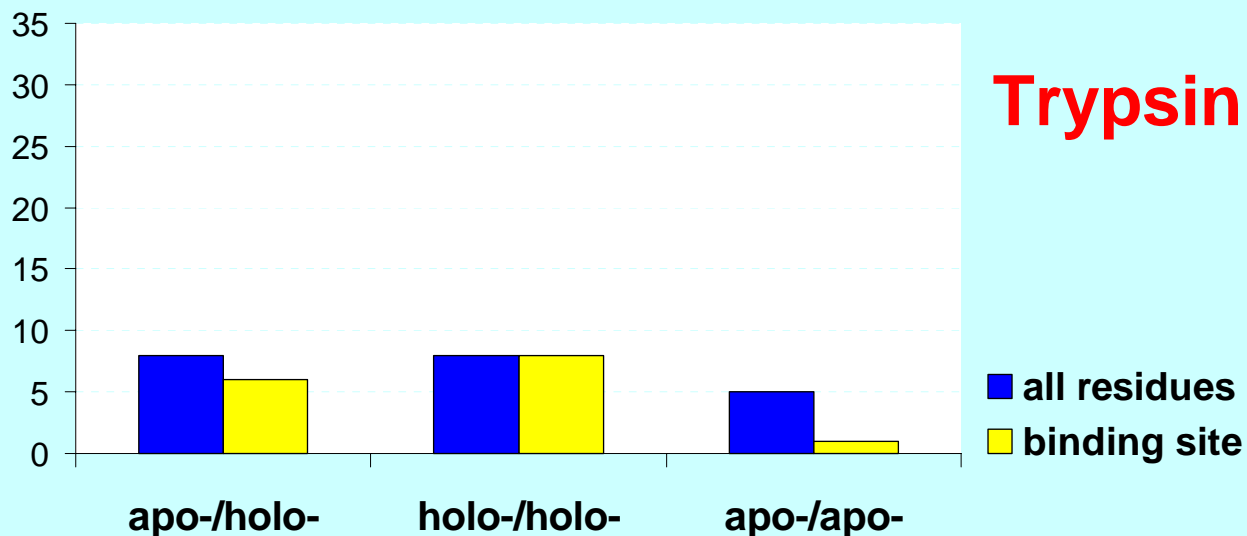
## 1. All environments conformational changes 1.1. Najmanovich threshold

Percentages of all side-chain torsions that change more than 60°



Resolution cut-off = 2.0 Å

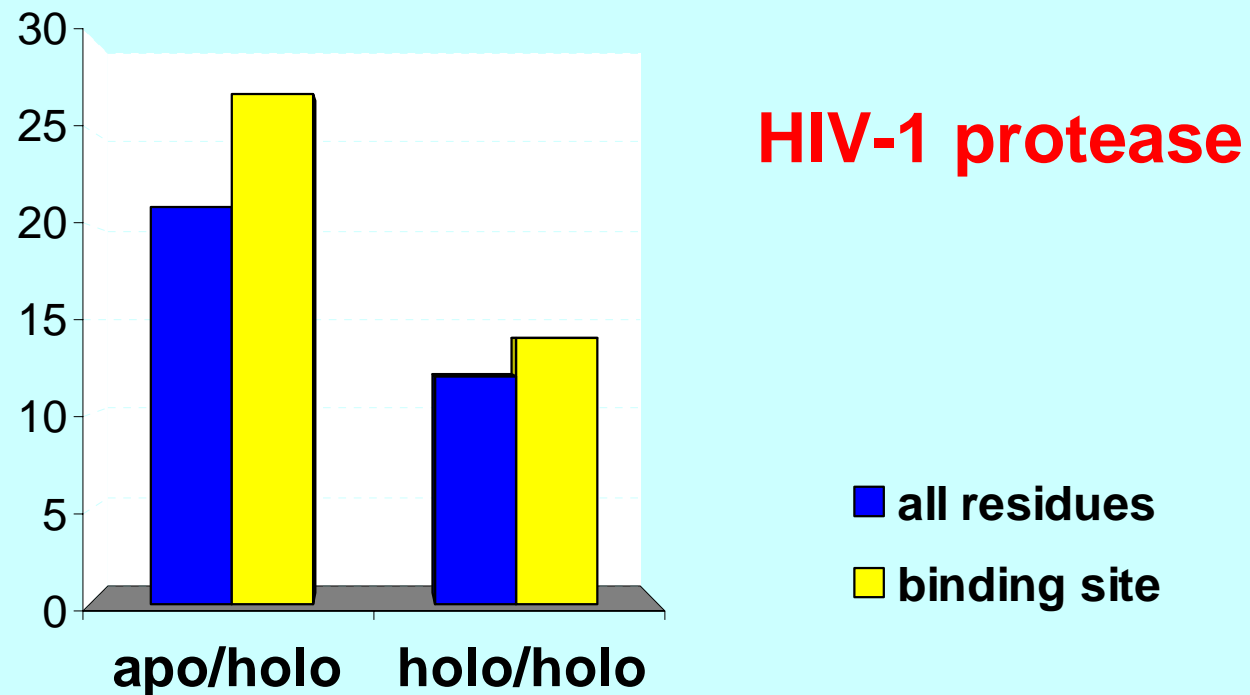
## Percentages of $\chi_1$ angles that change more than 60°



# Results

## 1.2. Rotamer libraries analysis

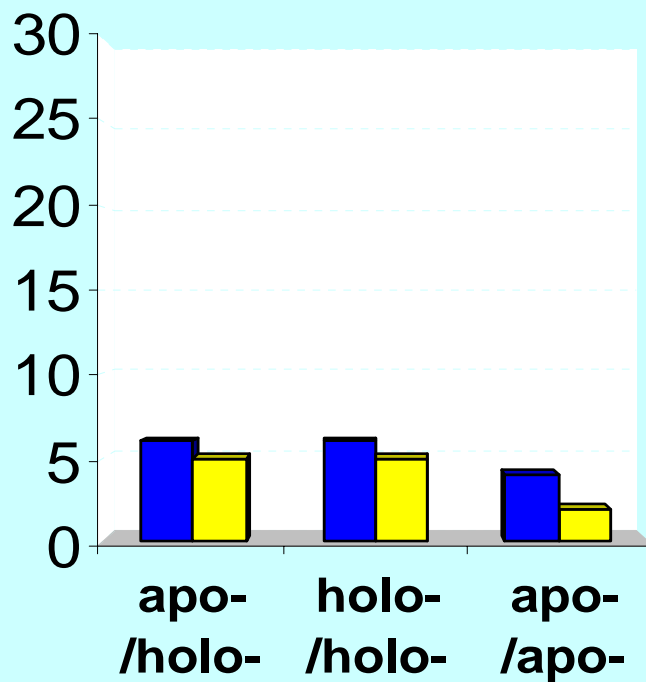
### Percentages of changed r1 rotamers



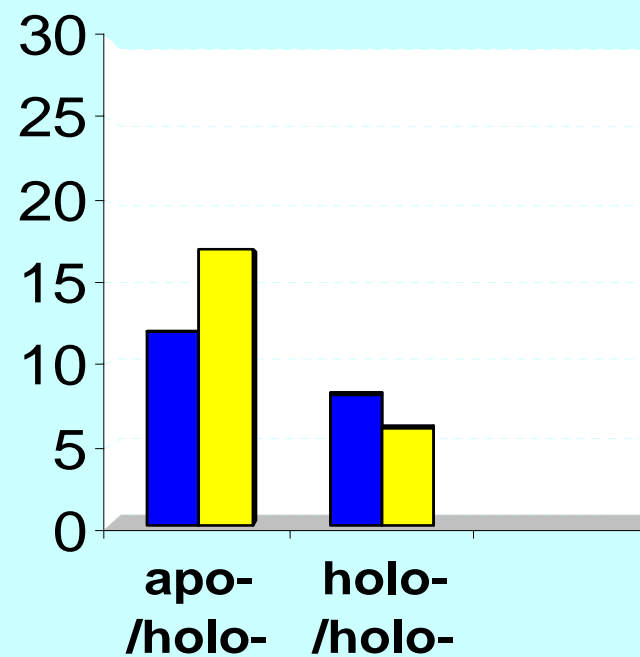
# Results

## Percentages of changed r1 rotamers

### Trypsin



### Thrombin



■ all residues ■ binding site

■ all residues ■ binding site

# Results

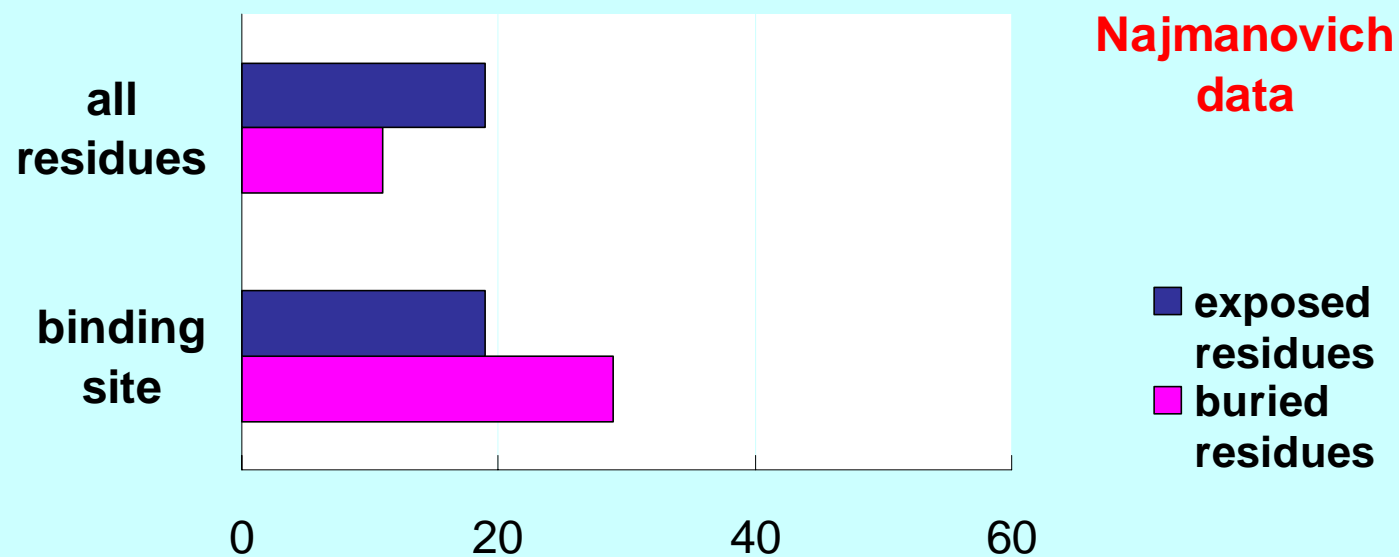
## 2. Differentiating buried and exposed residues

### 2.1. Najmanovich and Zhao thresholds

Percentages of  $\chi_1$  angles that change more than  $60^\circ$

#### HIV-1 Protease

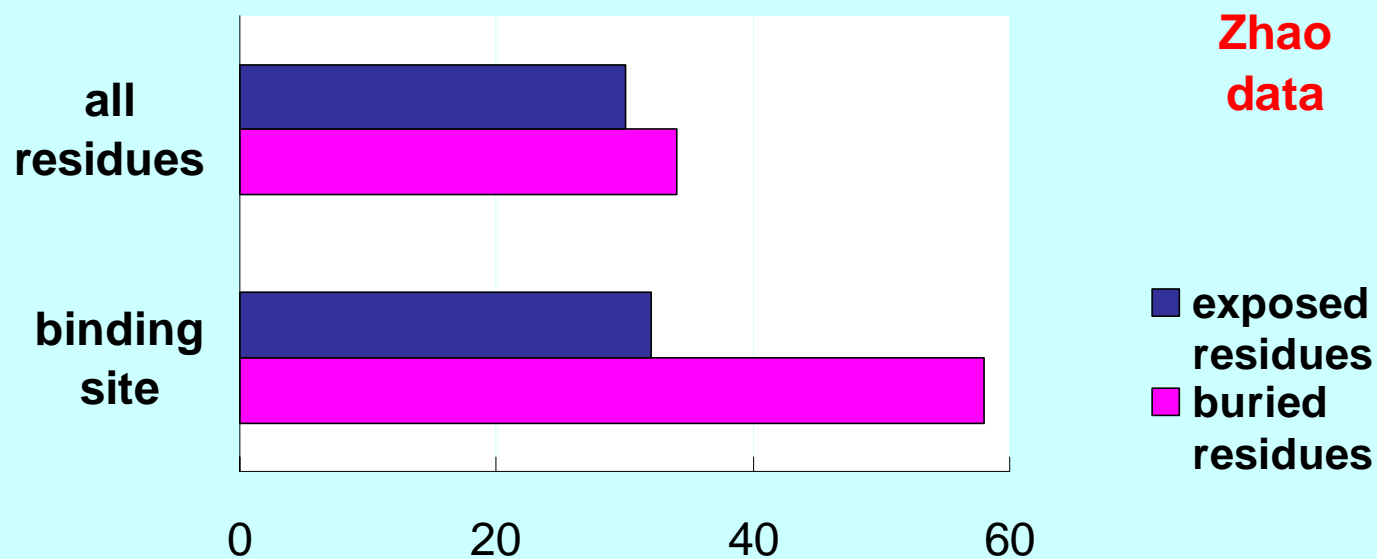
apo-/holo- comparisons



Percentages of  $\chi_1$  angles that change more than Zhao's environment- and residue type- specific thresholds

## HIV-1 Protease

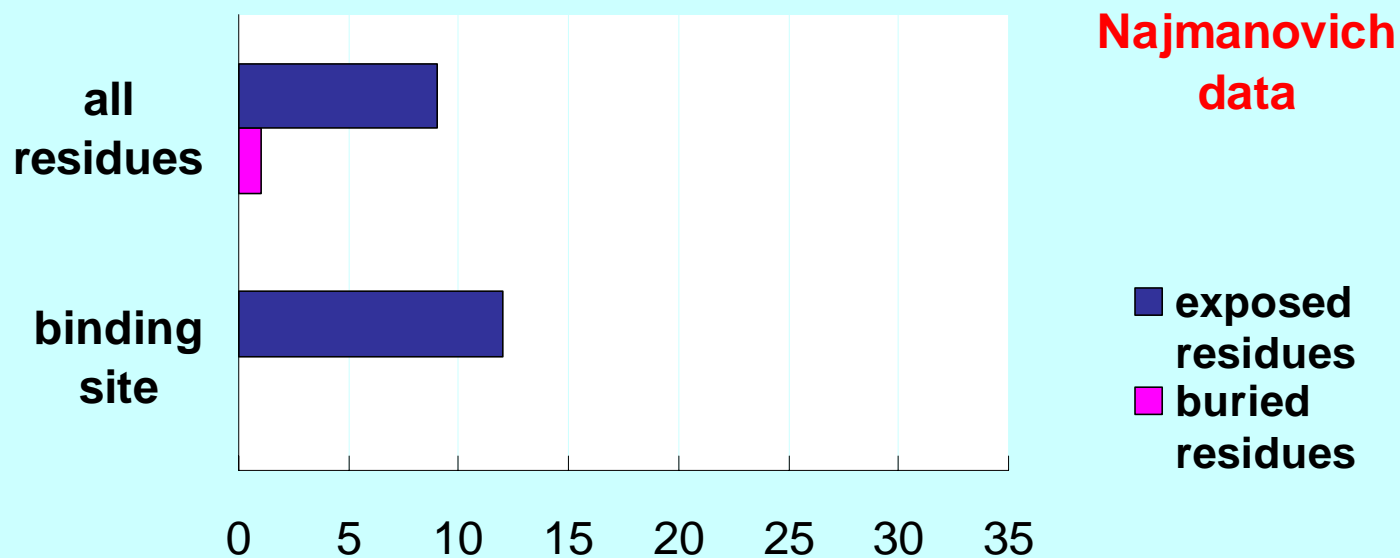
apo-/holo- comparisons



Percentages of  $\chi_1$  angles that change more than  $60^\circ$

**Trypsin**

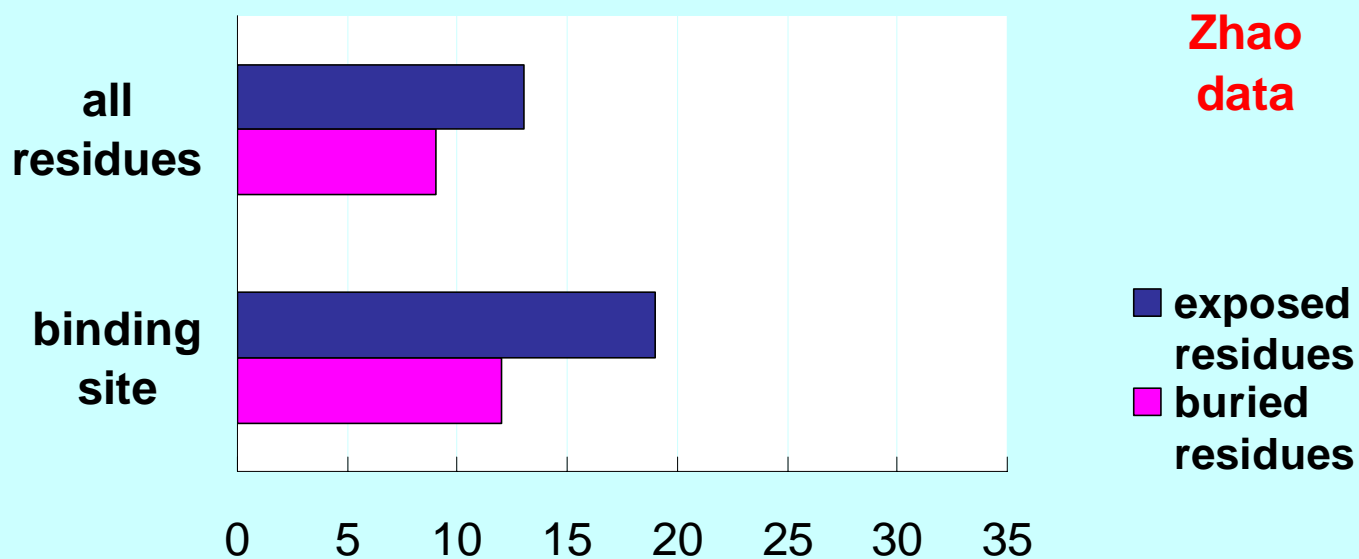
holo-/holo- comparisons



Percentages of  $\chi_1$  angles that change more than Zhao's environment- and residue type- specific thresholds

## Trypsin

holo-/holo- comparisons

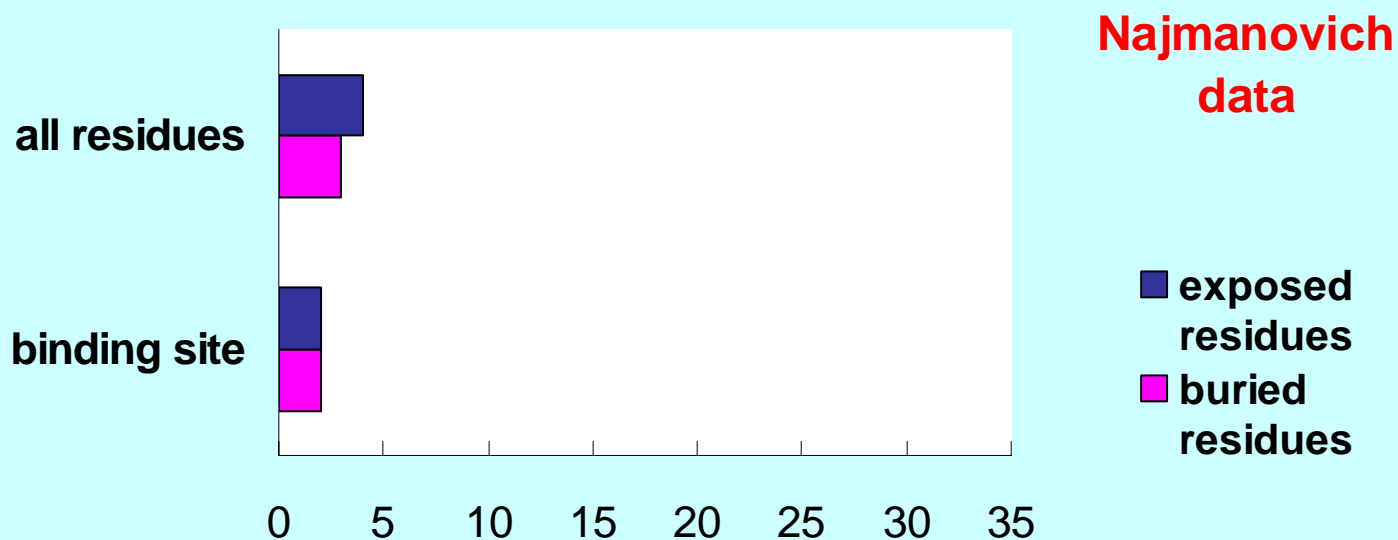


# Results

Percentages of  $\chi_1$  angles that change more than  $60^\circ$

## Thrombin

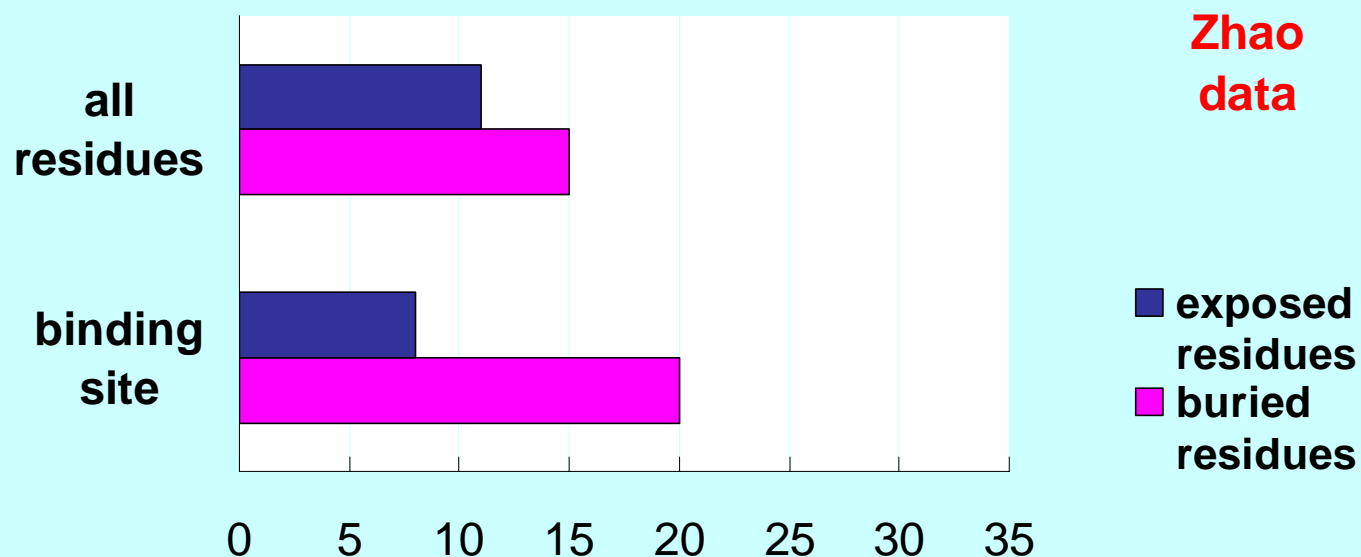
apo-/holo- comparisons



Percentages of  $\chi_1$  angles that change more than Zhao's environment- and residue type- specific thresholds

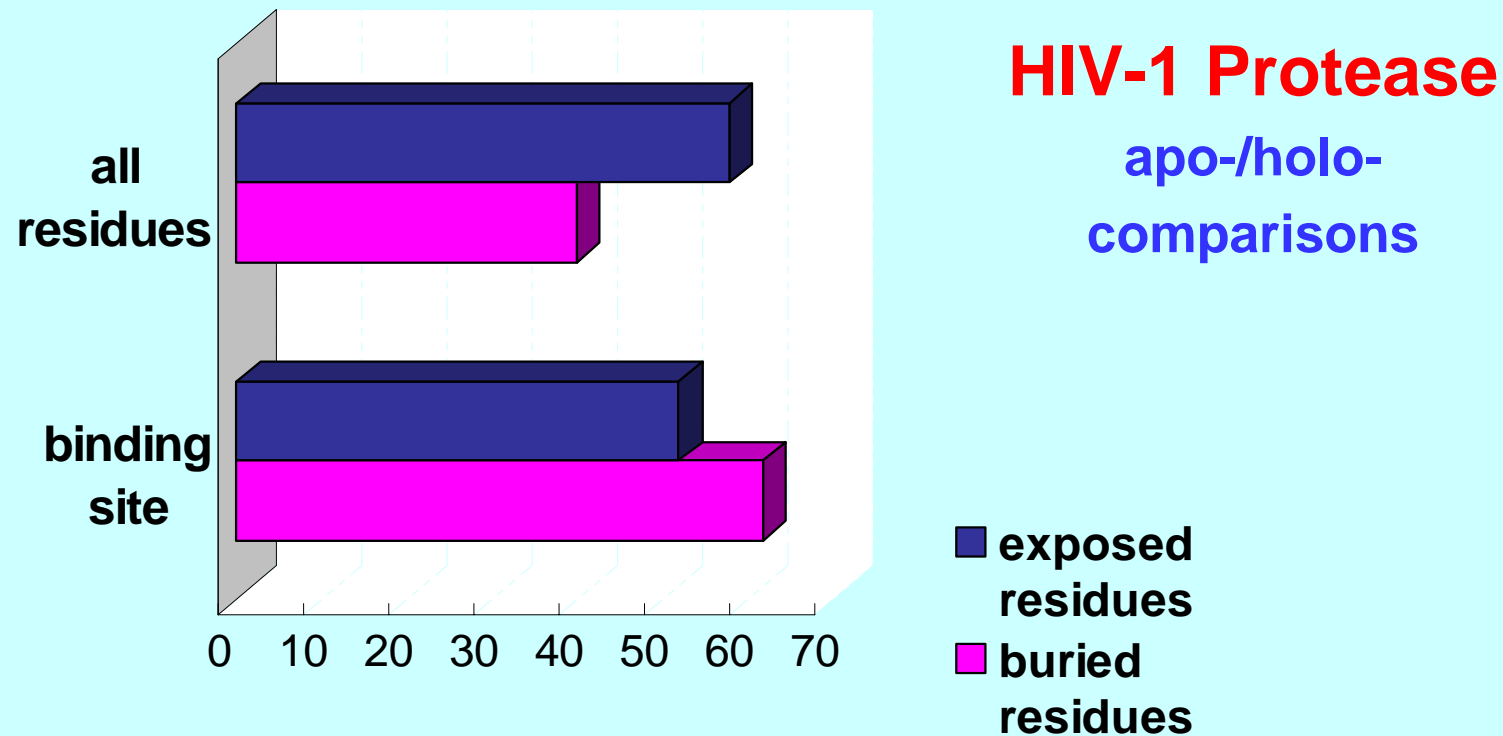
## Thrombin

apo-/holo- comparisons

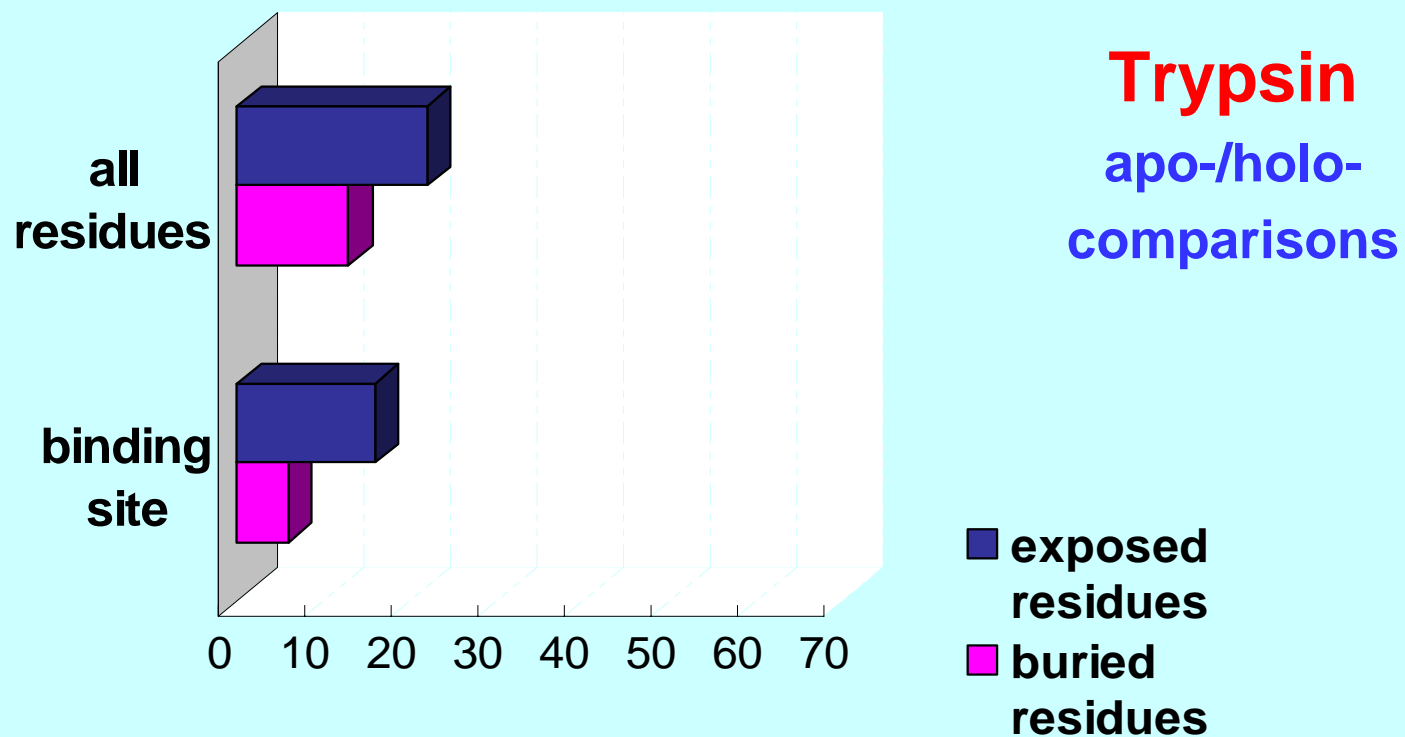


## 2.2. Rotamer library analysis

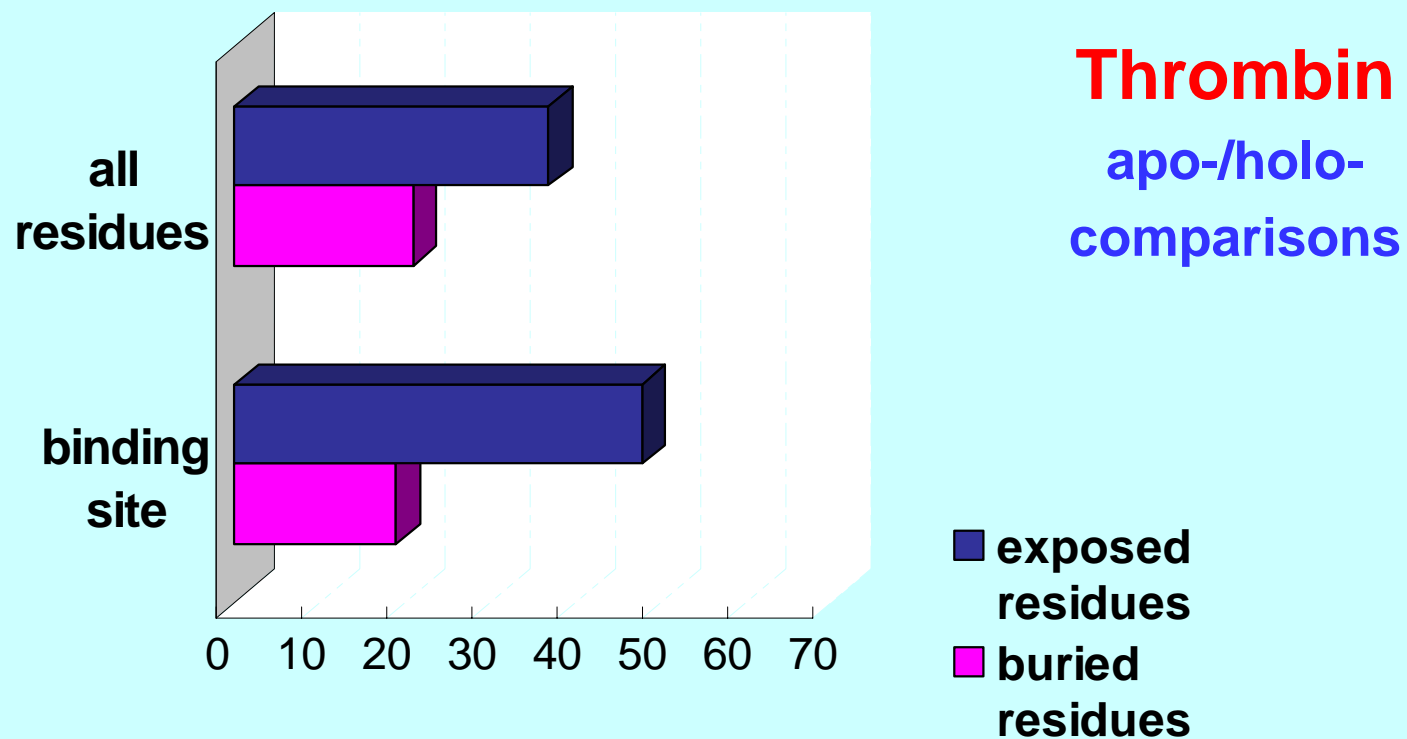
Percentages of buried and exposed residues that changed in *r1* and/or *r2* and/or rank



Percentages of buried and exposed residues that changed in *r1 and/or r2 and/or rank*



Percentages of buried and exposed residues that changed in *r1* and/or *r2* and/or rank



---

# Conclusions

---

- **Amino acid relative flexibilities:**
  - Broadly speaking in **agreement** with those reported in **literature**
  - Need to expand the dataset to avoid bias
- **Dissimilarity in the two chains** of **HIV-1 protease:**
  - **Asymmetries** in the mechanism of **binding**

# Conclusions

- **HIV-1 Protease:**

- **Flexibility** always greatest for **apo-/holo-** comparisons and **binding site** residues
- Probably a **systematic effect**

- **Trypsin and Thrombin:**

- **Smaller flexibility**
- **Trypsin binding site** residues **never more flexible** than all residues with **Najmanovich** and **Dunbrack** analyses
- **Dunbrack** analysis of **Thrombin** reveals **greater flexibility** in general, and greatest flexibility of **binding site** residues for **apo-/holo-** comparisons

# Conclusions

- **HIV-1 Protease:**

- **Buried residues** in the **binding site** **significantly more flexible** than exposed ones (**all methods**)

- **Trypsin:**

- **Exposed residues** are always the **most flexible** ones (**all methods**)

- **Thrombin:**

- **Buried residues** in the **binding site** **more flexible** than exposed ones applying **Zhao** thresholds

# Conclusions

- **Dunbrack probabilities:**
  - **Trends** can **not yet** be **generalised**
  - Need to enlarge the dataset

# Future (*and past!*) work

- **Dataset expansion**
- **ASA** (Accessible Surface Area) **variations** analysis
- **B-factors** analysis
- **Ligand/protein interactions** studies
- **Residue/residue** cross-correlated motions
- Use of **GRID**, **GOLPE**, **MIPSIM**

# Acknowledgments

Many thanks to:

Jonathan W. Essex, Andrew R. Leach and Paul A. Bamborough for their enthusiasm and advice

GlaxoSmithKline for funding this project