# The integration of tools for the preparation and analysis of PDB structures

Iuni M. L. Trist, Claudia Beato, Andrea Cristiani, Aldo Feriani, Alfonso Pozzan

Research Informatics – Computational Chemistry, Integrated Drug Discovery, Aptuit, an Evotec Company – Via A. Fleming 4, 37135 Verona (Italy)

## Introduction

3D PDB structures are a strong instrument for identifying key direct or water mediated protein-ligand interactions for a clear **understanding of the bonding mode** of compounds and support **rational drug design** efforts. They usually need to be carefully prepared to solve some issues (e.g. missing H atoms, alternate states, missing sidechains or loops, etc.). The MOE™,[1] Protonate 3D and Protein Preparation tools add H atoms and identify issues and suggest actions to take, but the **input of the user** is still very important for taking the final decisions, like giving the correct orientation to a particular chemical groups in the ligand, or removing water molecules not important for ligand binding. Simple visual inspection is not always enough and additional analyses (e.g. inspection of the electron-density map) are required. Moreover, while some types of intermolecular interactions (e.g. classical H-bonds and evident π-stack interactions) are well parameterized in molecular mechanics force fields, others are more difficult to identify because they are non-classical intermolecular forces (e.g. cation-π, CH-π, halogen-π, carbonyl n-π*, etc.), but still play fundamental roles. This is why the use of quantum mechanics (QM) is desirable for a more accurate understanding of ligand-protein interactions. Applications of QM methods in medicinal chemistry have been reported, and the Pair Interaction Energy Decomposition Analysis in particular (PIEDA) within the framework of the Fragment Orbital Method (FMO) has shown to be a powerful tool.[2,3] **Tools within MOE™** (e.g. 3D RISM solvent analysis, Energy Minimize with Optimal OH orientation) **as well as the integration with other software** (e.g. WinCoot4 for electron-density analysis or GAMESS[5] for QM approaches) **can give an important support.**

## PDB file preparation and important features of some protein-ligand complexes

- 8 PDB files prepared according to the methods below, by using MOE™ Protonate 3D and Protein preparation tools. Alternate residues were assigned according electron-density map analysis and termini were capped.
- **Water molecules** directly interacting with the ligand were observed in the preparation of 3 complexes (compounds **1**, **5** and **6**)
- Compound **2** has a terminal **amide group**. According to electron-density and hydrogen bond network analyses, no preferences on its **orientation** could be determined.
- For drug design purposes, the understanding of the involvement of interacting water molecules and of the orientation of the amide group compound **2** is very important
- **The integration of several tools and approaches were applied to determine the most reasonable system setup for these problematic PDB structures**

## Integration of tools for effective system preparation

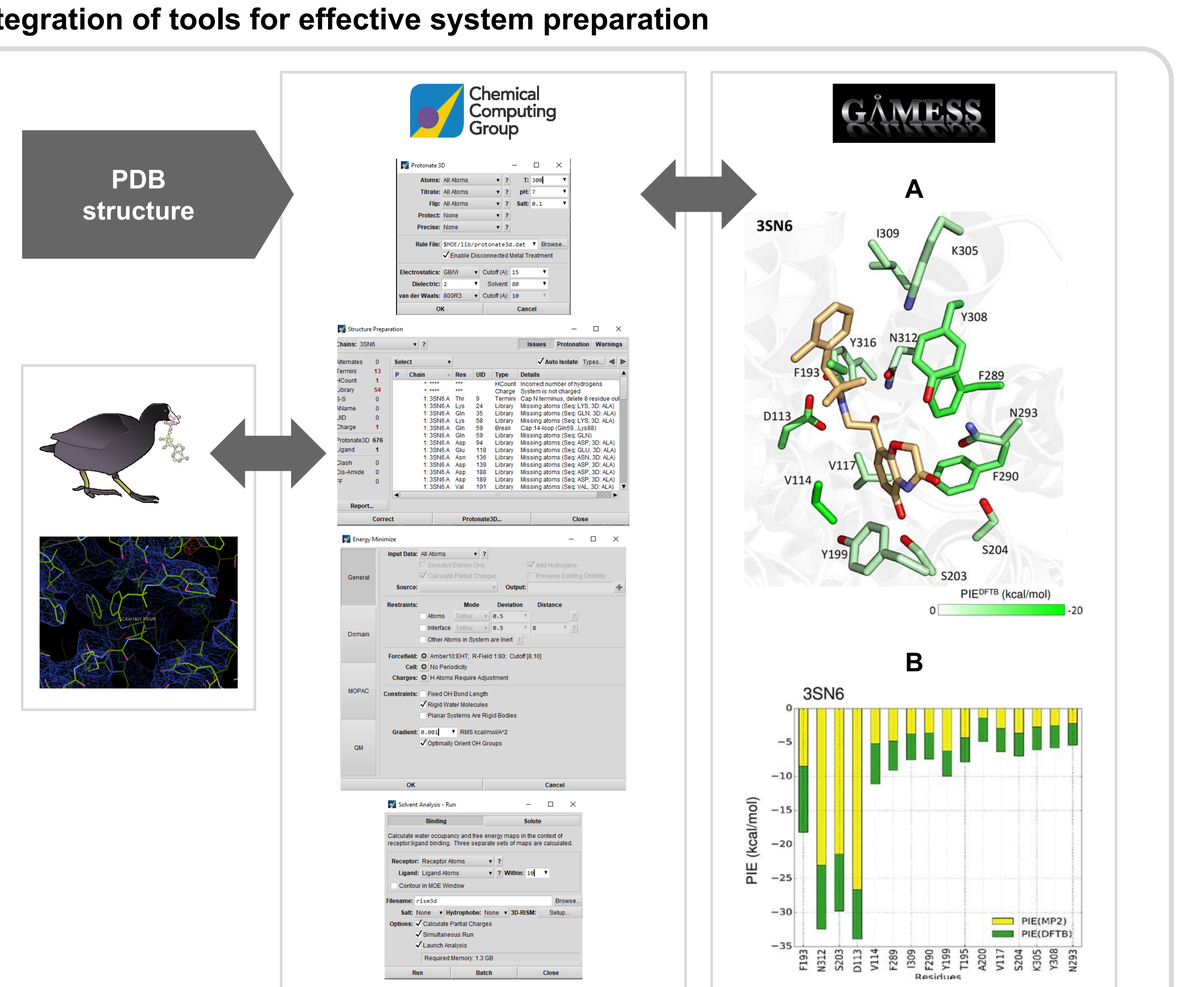


Figure 1: General approach for system preparation. Example and Figures in panels A and B were taken from Ref 5.

- Hydrogens were added to PDB structures using the MOE™ Protonate 3D tool[1], default settings
- The PDB structures were prepared using the MOE™ Protein preparation tool[1], default settings
- Added elements and H atoms were energy minimized using the MOE™ Energy Minimize tool[1], Gradient 0.001, OH groups optimally oriented
- PDB electron density was analyzed with Win Coot 0.8.6[4]
- Water map analysis was performed with MOE™ 3D RISM tool[1], default settings
- FMO input files were prepared using MOE™,[1] and a modified version of the FMOutil available on the SVL Exchange site[6]
- FMO-PIEDA calculations were performed with GAMESS[5]

## Relevant water molecules within the binding site

- **3D RISM** water map analysis classified the water molecules interacting with the ligand as **"happy"** in two cases (compounds 1 and 6) and as **"unhappy"** for compound **5**
- **FMO-PIEDA** calculations were performed for all the systems. Complexes with compounds **1**, **5** and **6**, calculations were prepared with and without the ligand interacting water molecule to address its importance in molecule binding.
- **Good correlation** between the FMO-PIEDA Totals and the experimental activity (pKi) when retaining the water molecules in compounds **1** and **6** complexes and removing the water molecule in the complex involving compound **5**, with **$R^2$** of **0.87**.
- **In line with 3D RISM analysis results**, a loss in predicted binding energy was registered by removing the water ligand interacting molecule from complexes **1** and **6** or by introducing the ligand interacting molecule in the complex involving compound **5**
- According to these results, the interacting water molecule was retained for complexes involving compounds **1** and **6**, while it was removed in compound **5** complex

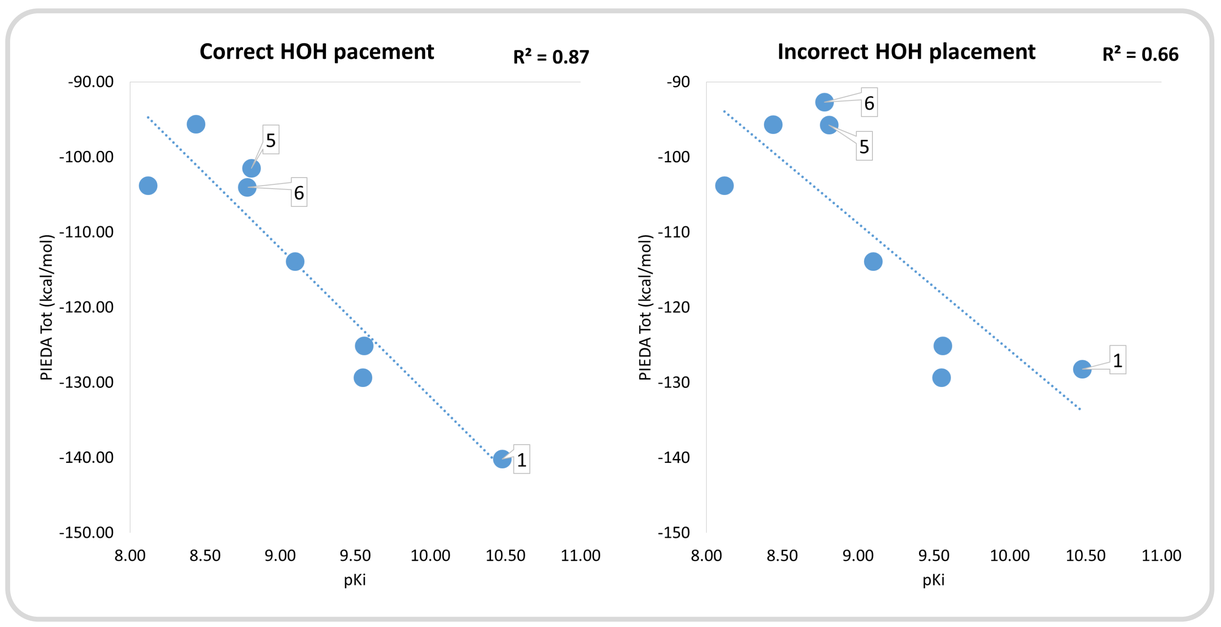


Figure 2: Scatter plot showing the experimental activity (pKi) against the predicted binding enthalpy (PIEDA Tot). Left: correct inclusion of the ligand-bound water molecule. Right: incorrect inclusion of the ligand-bound water molecule.

## Identification of the right conformation for ligand chemical groups

- Compound **2** has an external amide group
- FMO-PIEDA calculations were performed for all the systems. For the system involving compound **2**, the simulation was performed for the two possible made orientations.
- The **FMO-PIEDA** approach was **able to discriminate** between the two amide conformation, suggesting that one produces a stronger ligand binding
- A loss of correlation between the FMO-PIEDA Totals and the experimental activity (pKi) was observed for the other amide conformations
- According to these results, the most stabilizing amide conformation was retained for following drug design efforts

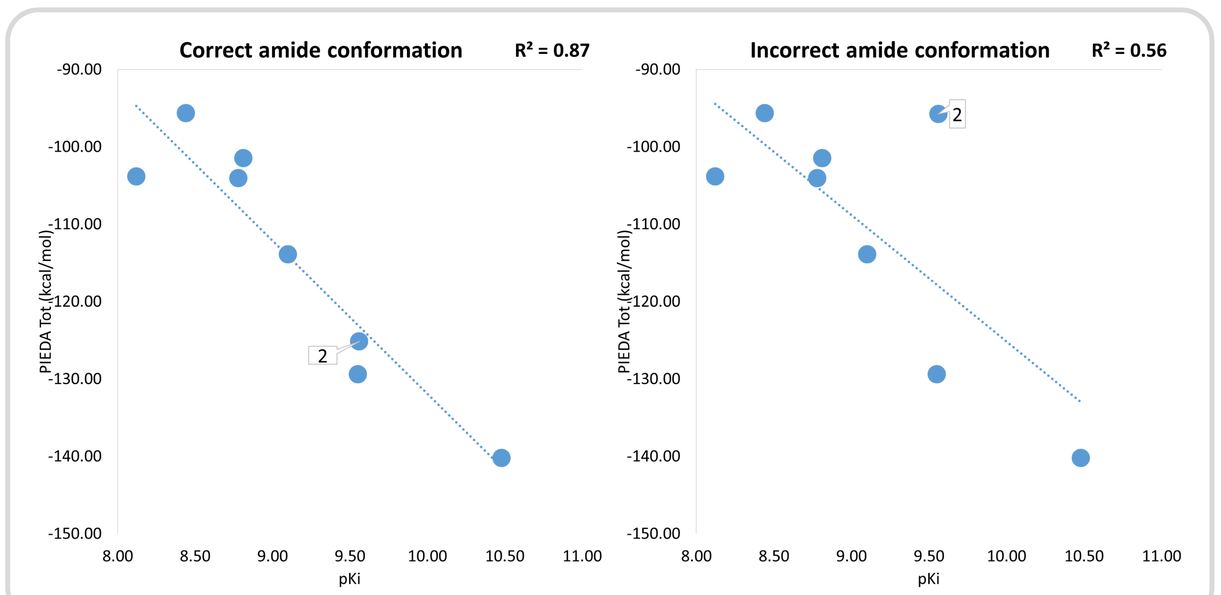


Figure 3: Scatter plot showing the experimental activity (pKi) against the predicted binding enthalpy (PIEDA Tot). Left: correct amide conformation for compound **2**. Right: incorrect amide conformation for compound **2**.

## Conclusions

- Appropriate preparation of PDB structures is pivotal for effective drug design efforts
- The use of only one tool is often not sufficient to correctly address all of the issues
- **The integrated use of tools within MOETM and from external software can give an important contribution for a correct PDB preparation and protein-ligand interaction analysis**

## Acknowledgments

The authors would like to thank the Computational Chemistry and the entire project teams in for fruitful discussions and support.

## References

1. Chemical Computing Group ULC, 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2018.
2. a) Fedorov, D.G., Kitaura, K., *J. Comp. Chem.,* 2006, 28, 222. b) Kitaura, K., Ikeo, E. et al., *J. Chem. Phys. Lett.,* 1999, 313, 701. c) Fedorov, D.G., Kitaura, K., *J. Phys. Chem.,* 2007, 111, 6904.
3. a) Morao, I. et al., *J Com Chem,* 2017, 38, 1987–90. b) Heifetz, A. et al., *J Med Chem,* 2016, 59, 4352−63.
4. Emsley, P. et al., *Acta Crystallogr D Biol Crystallogr,* 2010, 66, 486-501.
5. a) Schmidt, M. W. et al., *J ComputChem,* 1993, 14, 1347-1363. b) Gordon, M. S., C. E. Dykstra, G. Frenking, K. S. Kim, G. E. Scuseria, editors, Elsevier, Amsterdam, 2005.
6. FMOutil svl by Shinya Nakamura