Protein structure determination using NMR backbone chemical shifts

Marius Wanko[1], Christian Damblon[1]

[1]University of Liege, Department of chemistry, Structural Biological Chemistry Unit , Liege, 4000, Belgium

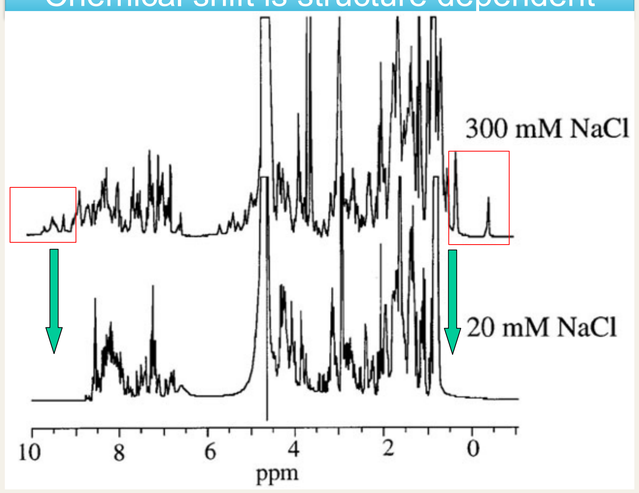
Introduction

The knowledge of the tridimensional structure of a protein is essential to design drugs, to predict protein function and to study mechanism of protein function. 3D structure can be determined using 2 experimental techniques: X-Ray and NMR. However, these techniques have limitations: they are time consuming, manually intensive and sometime technically difficult. Due to these limitations, different *In Silico* methods such as CS-HM-ROSETTA which combines NMR experimental data such as chemical shift (CS) that are structure dependent and homologous structure information have been employed for predicting 3D structures. CS-HM-ROSETTA is a fragment-based approach.

Chemical shift is structure dependent

Ribonuclease
Structured

Ribonuclease
Unstructured



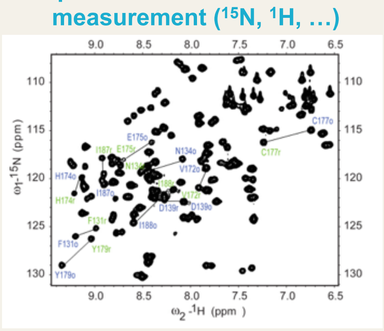
Objectif: Rapid protein structure determination using experimental NMR backbone chemical shift which are easily and quickly determined (at most 2 weeks) for predicting protein 3D structure.

CS-HM-Rosetta procedure summary

Protein sequence

MPFSFNLSSGNYLSTQDVEVLQ
RATRDHQMERLTIGERSFSVRY
QSAMDAFIVDPVQGELYSGLSHT
ELADIIRLADSVENQLNGTEGNL
GGWCHTQRKMRKQGKLYNDRR
LLLDKIGFVWSLEHHMNQNLQG
EWMKNYEELKS

Experimental Chemical Shifts measurement (^{15}N, ^{1}H, …)



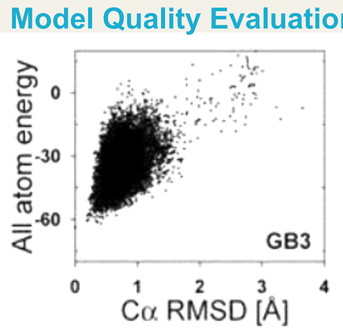
Fragments Selection

Fragments Assembly

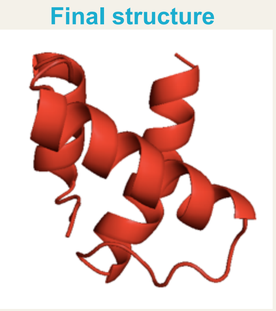
Model Quality Evaluation

Model Quality Evaluation

For each model, all atom energy is plotted against average distance between the model Cα atoms coordinates and the lowest energy model Cα atoms coordinates (RMSD). Therefore, the most probable model will correspond to the lowest energy and the lowest RMSD.

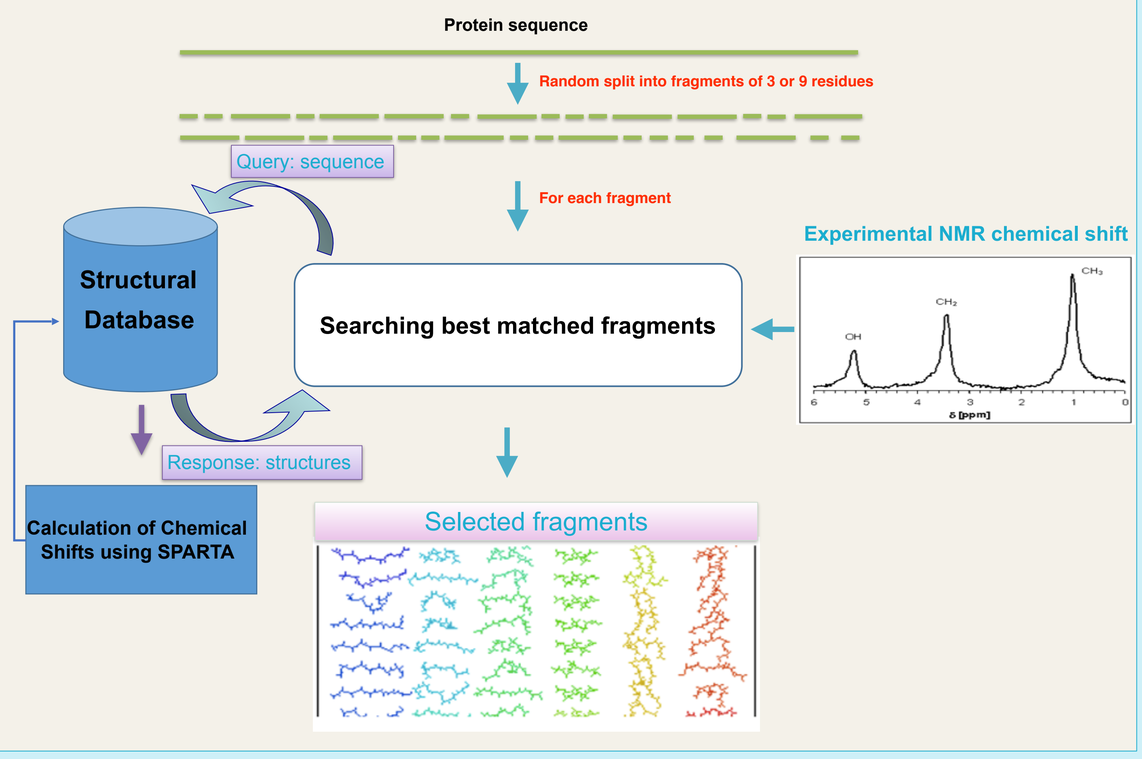


Final structure



Fragments Selection

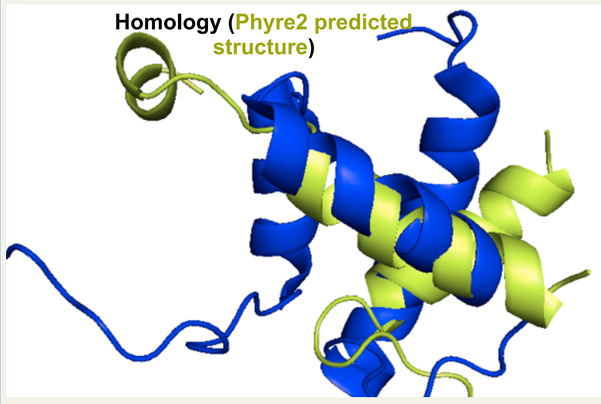
It is the first step of structure determination. Fragments are directly excised from protein database. For each 3 or 9 amino acids fragments, a library of 200 amino acids fragments were selected from protein database.



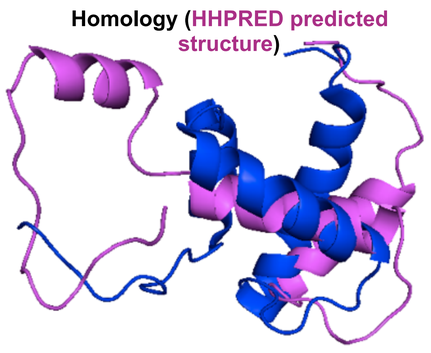
Results

Superimposition of experimental NMR structure of a domain of protein A6KY75 (blue) and predicted structures

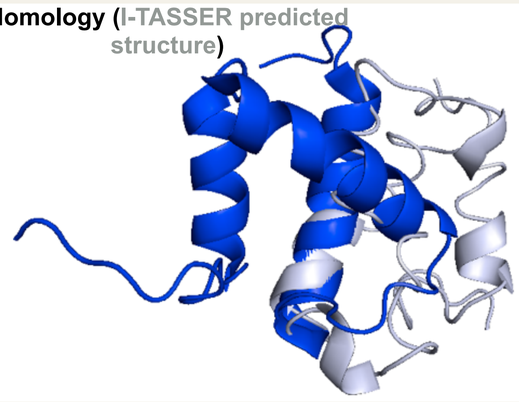
Homology (Phyre2 predicted structure)



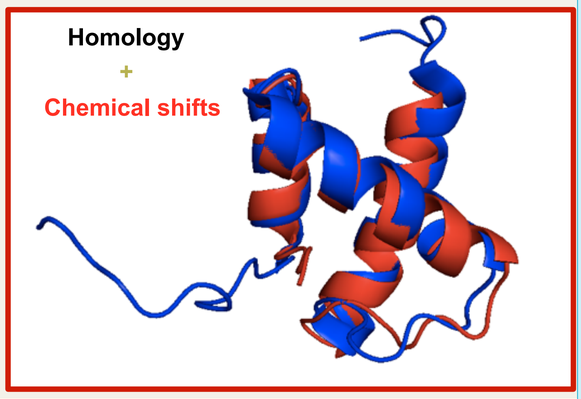
Homology (HHPRED predicted structure)



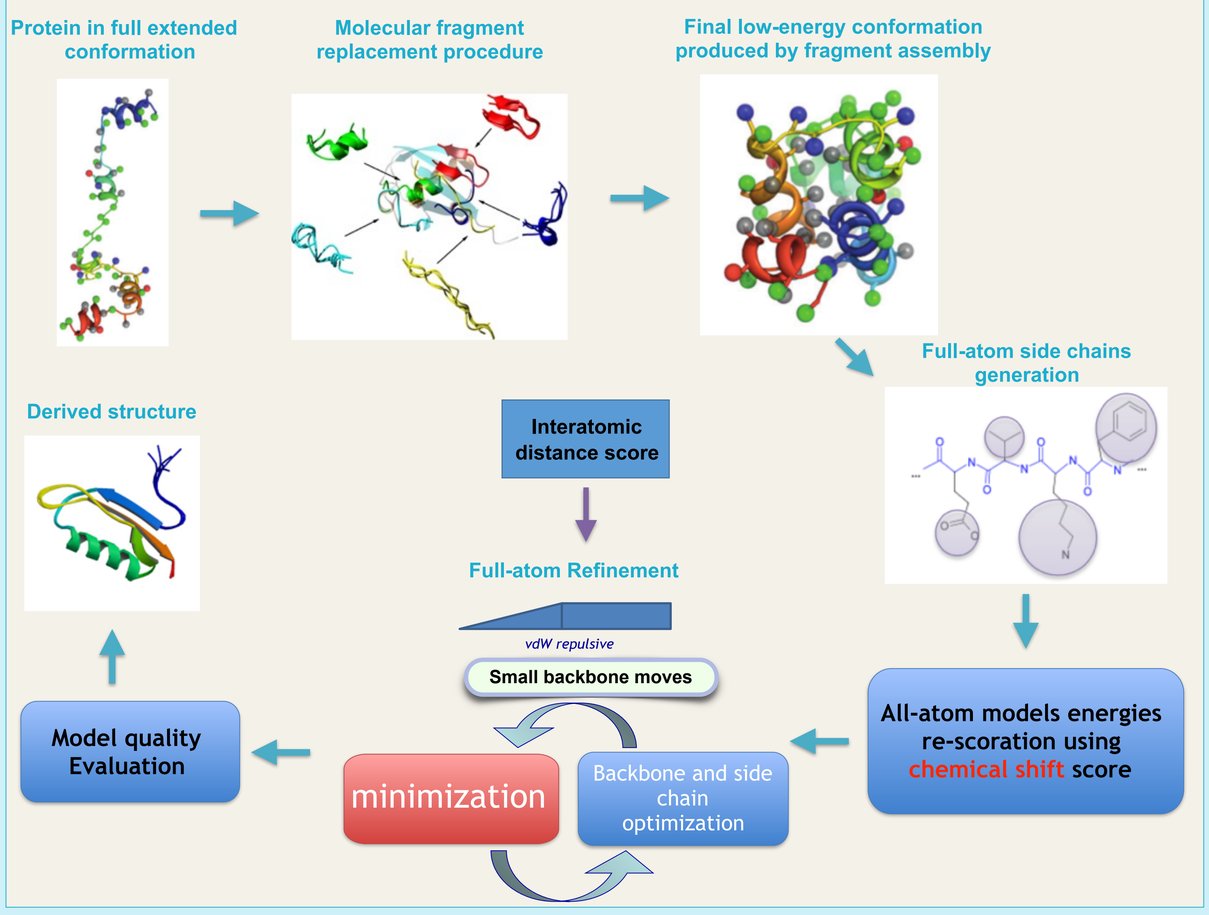
Homology (I-TASSER predicted structure)



Homology
+
Chemical shifts



Fragments Assembly



Conclusions

By using NMR chemical shifts, it is possible to determine reasonably accurate models of proteins. **As seen on the results, derived structures should be, more accurate than those determined using comparative modeling.**
Despite the fact that CS-HM-Rosetta method is a significant improvement over chemical shift based structure determination, improvements can be made. Thus, the main purpose of my thesis is the **development of a platform which uses simultaneously three chemical shift based methods to determine protein 3D structures.** This purpose has been divided into two goals:

1. Optimization of the **model quality evaluation** procedure
2. Evaluation of the **three existing methods** on large number of proteins whose structure was determined by NMR in order to study the complementarity between these methods
3. Development of a **platform** that will allow the rapid protein 3D structure determination using simultaneously these three methods and to determine the most probable structure using our optimization evaluation procedure

References

Andrea Cavalli, Xavier Salvatella, Christopher M. Dobson, and Michele Vendruscolo. PNAS. **2007**, 104, 9615 .

Yang Shen, Oliver Lange, Frank Delaglio, Paolo Rossi, James M. Aramini, Gaohua Liu, Alexander Eletsky, Yibing Wu, Kiran K. Singarapu, Alexander Lemak, Alexandr Ignatchenko, Cheryl H. Arrowsmith, Thomas Szyperski, Gaetano T. Montelione, David Baker, Ad Bax. PNAS. **2008**, 105, 4685.

David S. Wishart, David Arndt, Mark Berjanskii, Peter Tang, Jianjun Zhou, Guohui Lin. Nucleic Acids Research. **2008**, 36, W496

Yang Shen, Robert Vernon, David Baker, Ad Bax. J Biomol NMR. **2009**, 43, 63

James M. Thompson, David Baker. Proteins. **2011**, 79, 2380.

James M. Thompson, Nikolaos G. Sgourakis, Gaohua Liu, Paolo Rossi, Yuefeng Tang, Jeffrey L. Mills, Thomas Szyperski, Gaetano T. Montelione, David Baker. PNAS. **2012**, 109, 9875.