

## COMPUTATIONAL METHODS FOR PROTEIN MISFOLDING IN ALZHEIMER'S DISEASE

Avramouli Antigoni<sup>1\*</sup>, Polichronidou Eleutheria<sup>1</sup>, Psiha Maria<sup>1</sup><sup>1</sup> Bioinformatics and Human Electrophysiology Laboratory, Department of Informatics, Ionian University, Corfu, Greece,

**Introduction :** The majority of protein molecules must be stable and folded into defined three dimensional structures to perform their functional activity. However, cells usually contain correctly folded proteins together with a multitude of conformational states. These misfolded proteins are prone to forming toxic aggregates, including soluble oligomers and fibrillar amyloid deposits, being the underlying cause of several severe and incurable age related disorders linked with neurodegeneration such as Alzheimer's disease [1]. To ensure protein homeostasis (or proteostasis) and avert protein aggregation, cells have developed a wide-ranging network including molecular chaperones and other factors [2]. These exquisite machineries degrade or either refold correctly misfolded proteins, but they tend to decay during aging, thus enabling the development of aggregate deposition diseases. Knowledge of the impact of protein folding has led to an explosion of work in protein structure determination and prediction, further facilitating drug design. Molecular experimental methods mostly used for the determination of a protein's structure are time consuming and high-priced. Thus, a remarkably wide range of computational prediction methods that can precisely, rapidly and automatically categorize unknown protein sequences into definite fold classes is required. Bioinformatics and computational biology has made remarkable progress in recognition of protein structures and a variety of computational prediction methods has been generated.

**Alzheimer is a folding disease:** Alzheimer's disease (AD) is a neurodegenerative disorder and is characterized clinically by cognitive decline, memory loss, visuospatial and language impairment and the commonest form of dementia [3]. The disease is histopathologically confirmed at post-mortem within the brain by the presence of neurofibrillary tangles and amyloid plaques [4, 5]. Regions of the brain that are involved in short-term memory and learning such as the temporal and frontal lobes are impaired because of neuronal loss and the breakdown of the neuronal synaptic connections. The greatest risk factor for contracting the disease is advancing age and risk is significantly increased beyond the age of 70 years. Furthermore, AD is an irreversible progressive dementia with long prodromal stages and up to now there is a lack of effective pharmacotherapy options. During the past decade, intense research was concentrated in discovering disease modifying small molecules as therapeutic options. This led to the understanding of the importance of proper protein folding and aggregation, characterizing AD as a folding disease. While the clinical symptoms of the disease are defined by cognitive impairment, the causes leading to memory decline are strongly tied to deposits of misfolded protein aggregates. The amyloid beta (A $\beta$ ) and neurofibrillary tangles (NFTs) are the hallmark deposits in AD brains. These aggregates derive from extremely vital and naturally occurring protein structures in the brain [6]. Why aggregation is so toxic to cells, especially to neurons, is not completely understood [7]. Notably, the toxic effects of aggregation underlie common structural properties of the aggregates and may be completely different to the normal function of the affected protein.

**Protein fold recognition through computational methods:** The elucidation of how amino acid sequences lead to correctly folded and fully functional proteins in cells remains one of the greatest challenges in science. This understanding will have a significant effect in various fields of biology and medicine and will mostly lead to the rational design strategies to generate new pharmaceutical proteins and drug molecules. In consequence, the determination of the fold category of a protein, a process that is called fold recognition, is fundamental revealing the tertiary structure of proteins. Traditional experimental methods that are used for the determination of a protein structure are X-ray crystallography and nuclear magnetic resonance spectroscopy. Since the completion of the Human Genome Project, a vast number of protein sequences is rapidly generated by next-generation sequencing techniques. Although many of these sequences are structurally characterized using experimental methods, a cumulative gap is created between the structurally determined sequences and uncharacterized ones. Therefore, the development of computational methods for rapid and precise determination of protein folds is obligatory.

Consequently, many computational methods for the accurate prediction of protein structures have been recently developed offering an alternative approach to the demanding, time consuming and high-priced experimental methods [8]. Protein fold recognition is studied with computational methods that can be largely categorized into the following classes:

- 1) de novo modeling methods. This series of approaches require long computational time and numerous resources, while they can only be effectively applied in small proteins.
- 2) template-based methods which are used for the determination of protein structures comparing proteins that are evolutionary related. These methods are considered very effective for the construction of theoretical models of protein structures and
- 3) template-free methods, that are based mainly on amino acid sequences to build the model and predict a protein's structure precisely.

**Conclusion:** Constantly there is active research to define the mechanisms by which disease-associated proteins misfold and create aggregates causing cellular toxicity. These processes are the hallmark of many incurable pathologies such as Alzheimer's disease. Continued progress in our ability to determine precise protein structure will utilize rational design strategies that will result in discovering novel pharmaceutical molecules. Compared with the traditional experimental methods, computational methods present many advantages such as the demonstration of robust, accurate and consistent performance, they can be applied in large-scale protein fold recognition. Furthermore, they can efficiently address the intrinsic restrictions of experimental methods, that is, their being labor intensive and expensive.

**References :**

1. Hartl, F.U. (2017). Protein Misfolding Diseases. *Annu. Rev. Biochem.* 86:28.1-28.6
2. Brehme, M., et al. (2014). A chaperome subnetwork safeguards proteostasis in aging and neurodegenerative disease. *Cell Rep.* 9:1135-50
3. Selkoe, D.J. (2004) Cell biology of protein misfolding: the examples of Alzheimer's and Parkinson's diseases. *Nat Cell Biol.* 6(11):1054-61.
4. Chiti, F. and Dobson, C.M. (2006). Protein misfolding, functional amyloid, and human disease. *Ann Rev Biochem* 75, 333-366.
5. J. Ávila, F. Lim, F. Moreno, C. Belmonte, C. Cuervo. (2002). Tau function and dysfunction in neurons. *Mol. Neurobiol.* 25, 213-231
6. P. Sweeney, et al. (2017) Protein misfolding in neurodegenerative diseases: implications and strategies. *Transl Neurodegener.* 6
7. Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., et al. (2002). Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* 416, 507-511.
8. Wei, L. and Zou, Q. (2016). Recent Progress in Machine Learning-Based Methods for Protein Fold Recognition. *Int. J. Mol. Sci.* 17, 2118.

**Key words :** protein folding, protein misfolding, protein aggregation, Alzheimer's disease, protein homeostasis

*Avramouli Antigoni: Computational methods for protein misfolding in Alzheimer's Disease*

2017 Corfu Workshop  
on Medical Physics & Biomedical Engineering

**COMPUTATIONAL METHODS FOR  
PROTEIN MISFOLDING  
IN ALZHEIMER'S DISEASE**

Avramouli Antigoni,  
Molecular Biologist, PhD candidate, Ionian University



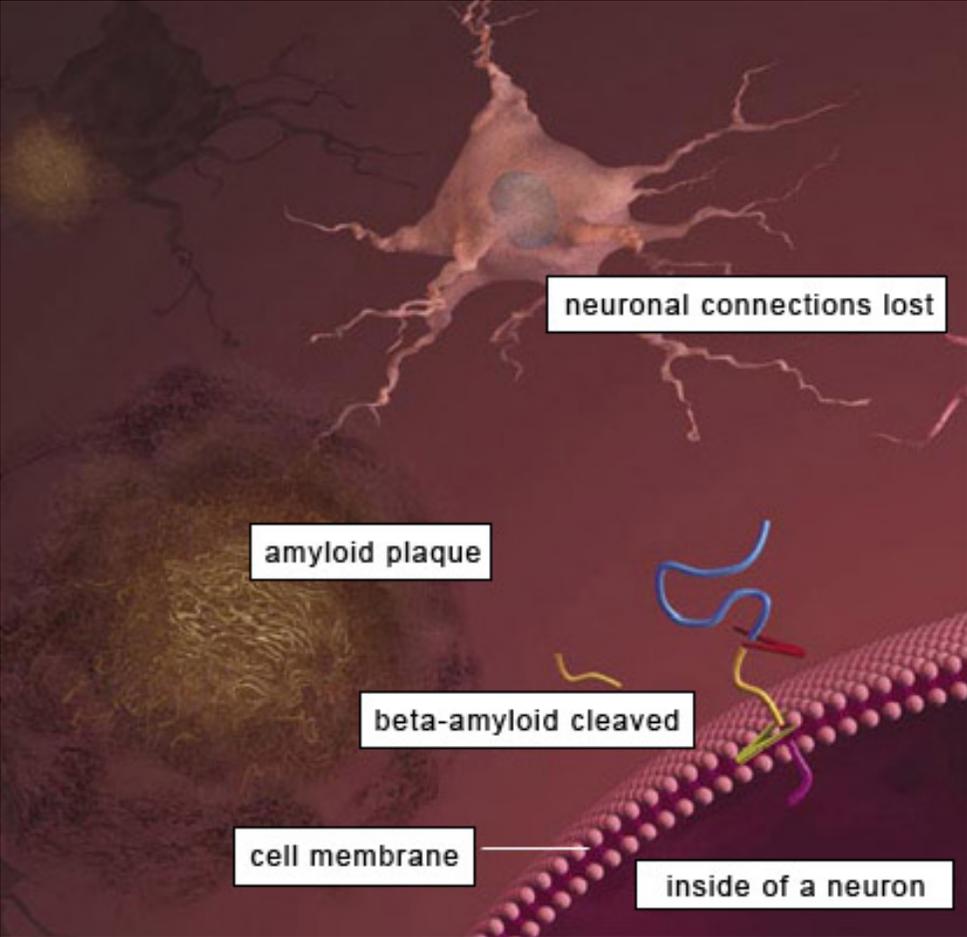
# Presentation outline

1. Proper protein folding
2. Computational methods for protein fold recognition
3. In silico study of the tertiary structure of well-established mutations in Alzheimer's Disease

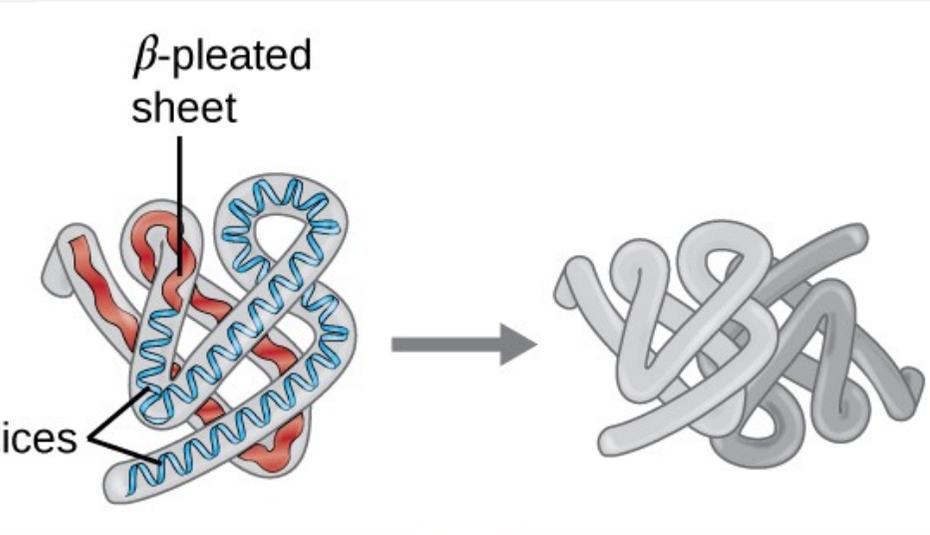
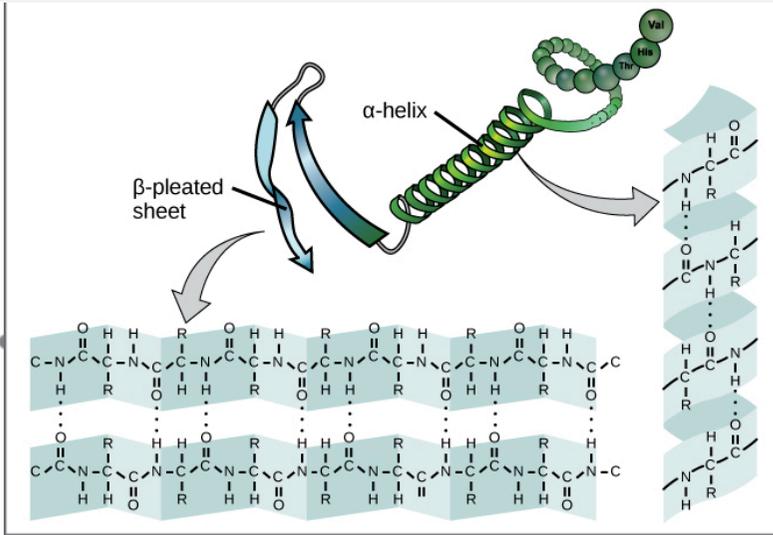
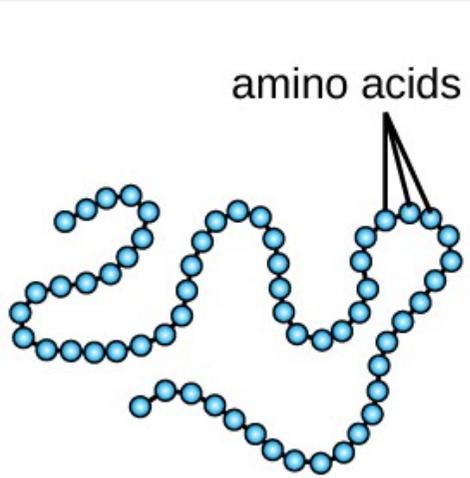


# Human diseases caused by defects in protein folding, stability and aggregation

Disease	Protein affected	Description
Cystic fibrosis	Cystic fibrosis transmembrane conductance regulator (CFTR)	The $\Delta$ Phe508 mutant has wild-type activity, but impaired folding in the endoplasmic reticulum leads to degradation.
$\alpha$ 1 Antitrypsin deficiency	$\alpha$ 1 Antitrypsin (also known as SERPINA1)	80% of Glu342Lys mutants misfold and are degraded. Pathology is due to aggregation in patients with a reduced degradation rate.
SCAD deficiency	Short-chain acyl-CoA dehydrogenase (SCAD)	Impaired folding of Arg22Trp mutants leads to rapid degradation.
Alzheimer disease	Presenilin, $\gamma$ -secretase	Mutations cause incorrect cleavage by the $\gamma$ -secretase protease to produce the amyloid $\beta$ -peptide; this aggregates into extracellular amyloid plaques.
Parkinson disease	$\alpha$ -Synuclein	Oxidative damage causes misfolding and aggregation. Hereditary forms are linked to deficiency in ubiquitin-mediated degradation.
Huntington disease	Huntingtin	CAG expansions in the Huntingtin gene lead to an abundance of polyglutamine fragments that aggregate and associate non-specifically with other cellular proteins.
Sickle cell anaemia	Haemoglobin	The Glu6Val mutation leads to aggregation in red blood cells.







**Primary Protein Structure**  
Sequence of a chain of amino acids

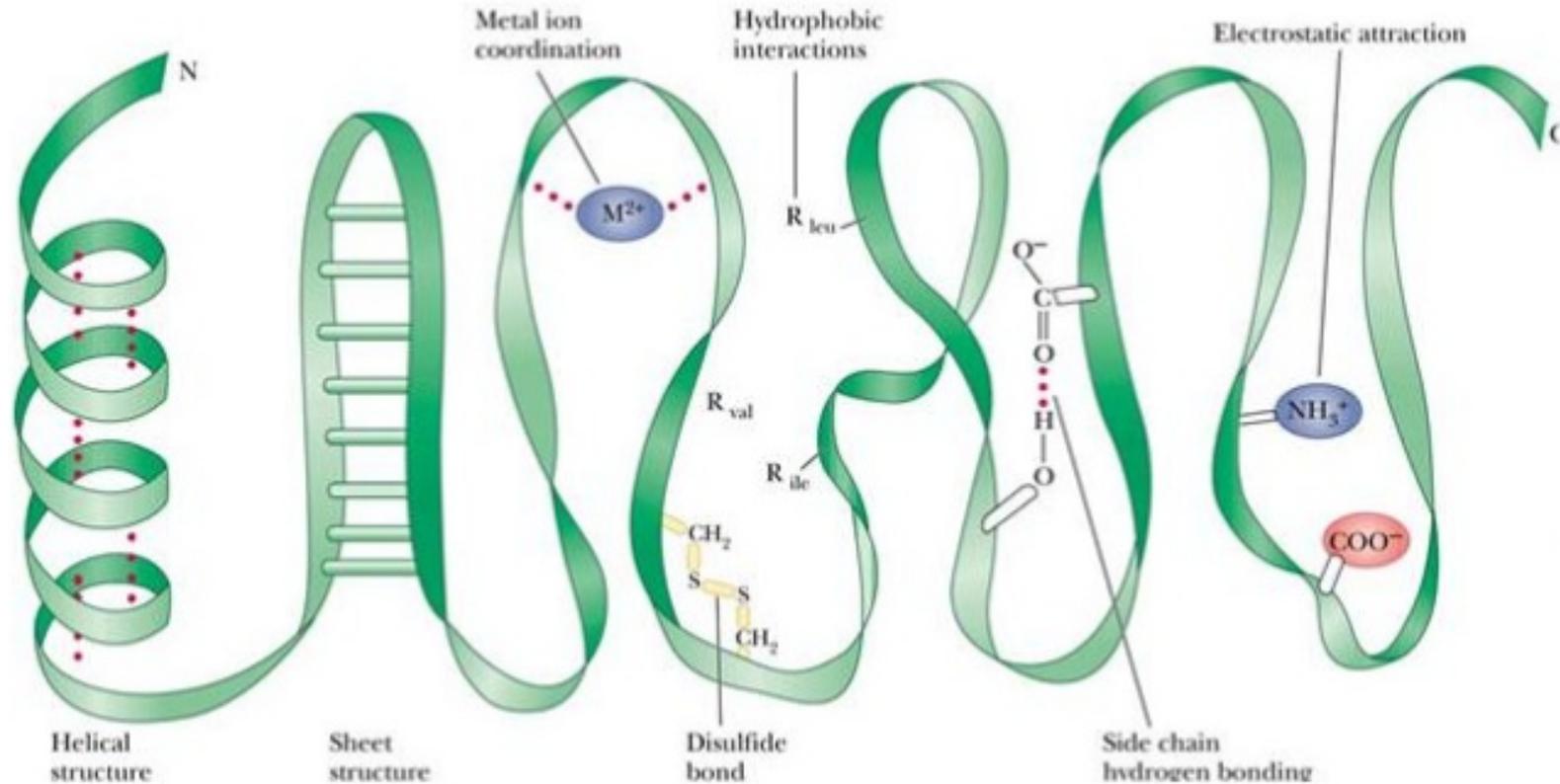
**Secondary Protein Structure**  
Local folding of the polypeptide chain into helices or sheets

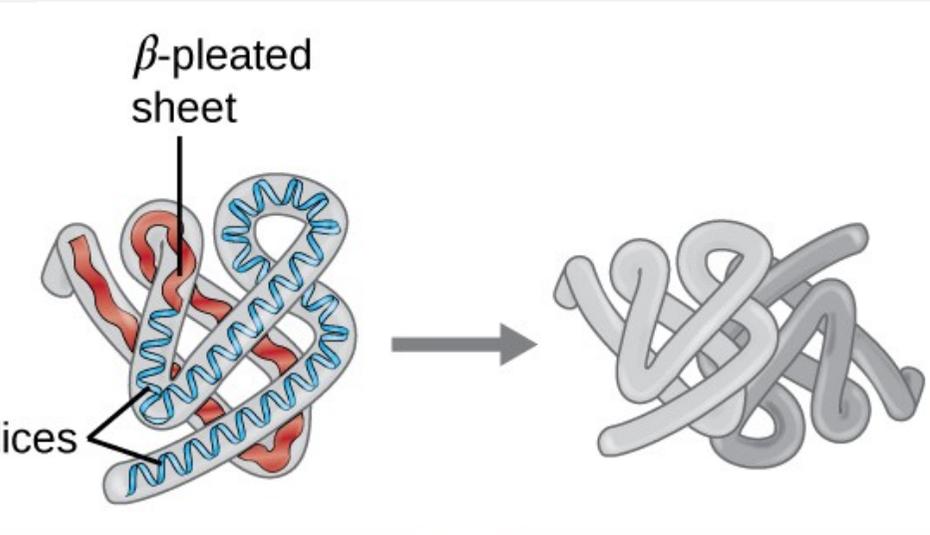
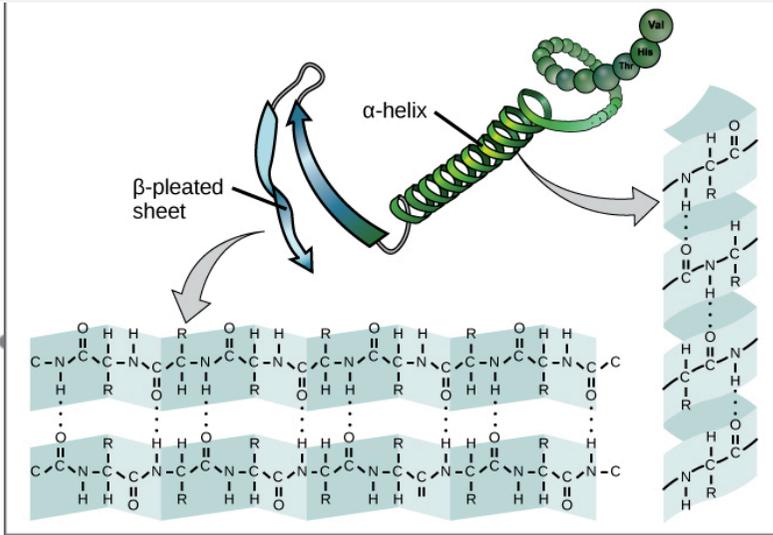
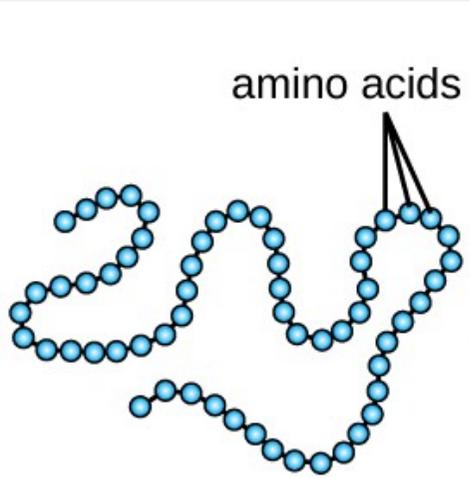
**Tertiary Protein Structure**  
three-dimensional folding pattern of a protein due to side chain interactions

**Quaternary Protein Structure**  
protein consisting of more than one amino acid chain

# Forces that stabilize protein structure

- Interactions between atoms within the protein chain
- Interactions between the protein and the solvent





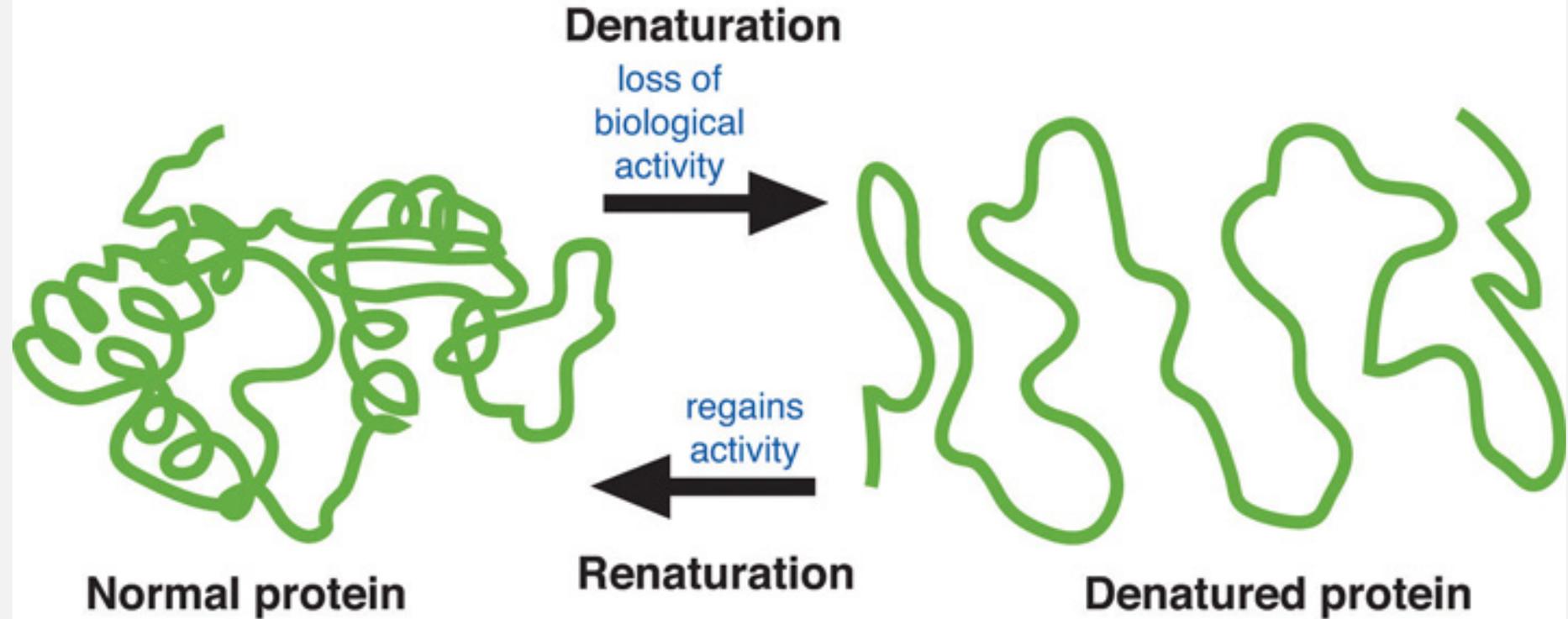
**Primary Protein Structure**  
Sequence of a chain of amino acids

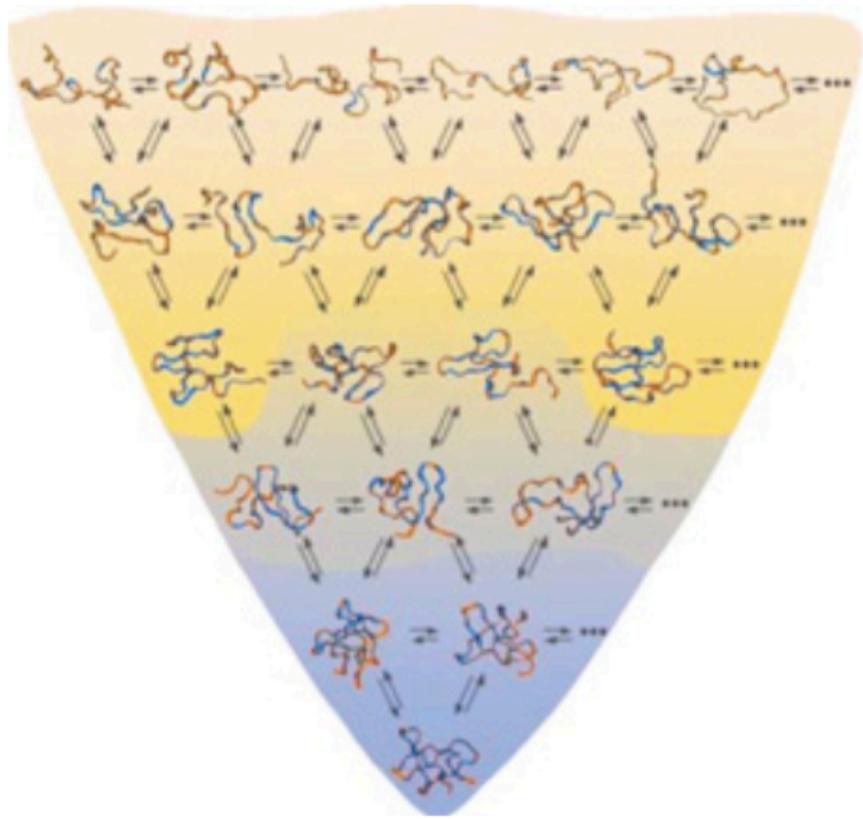
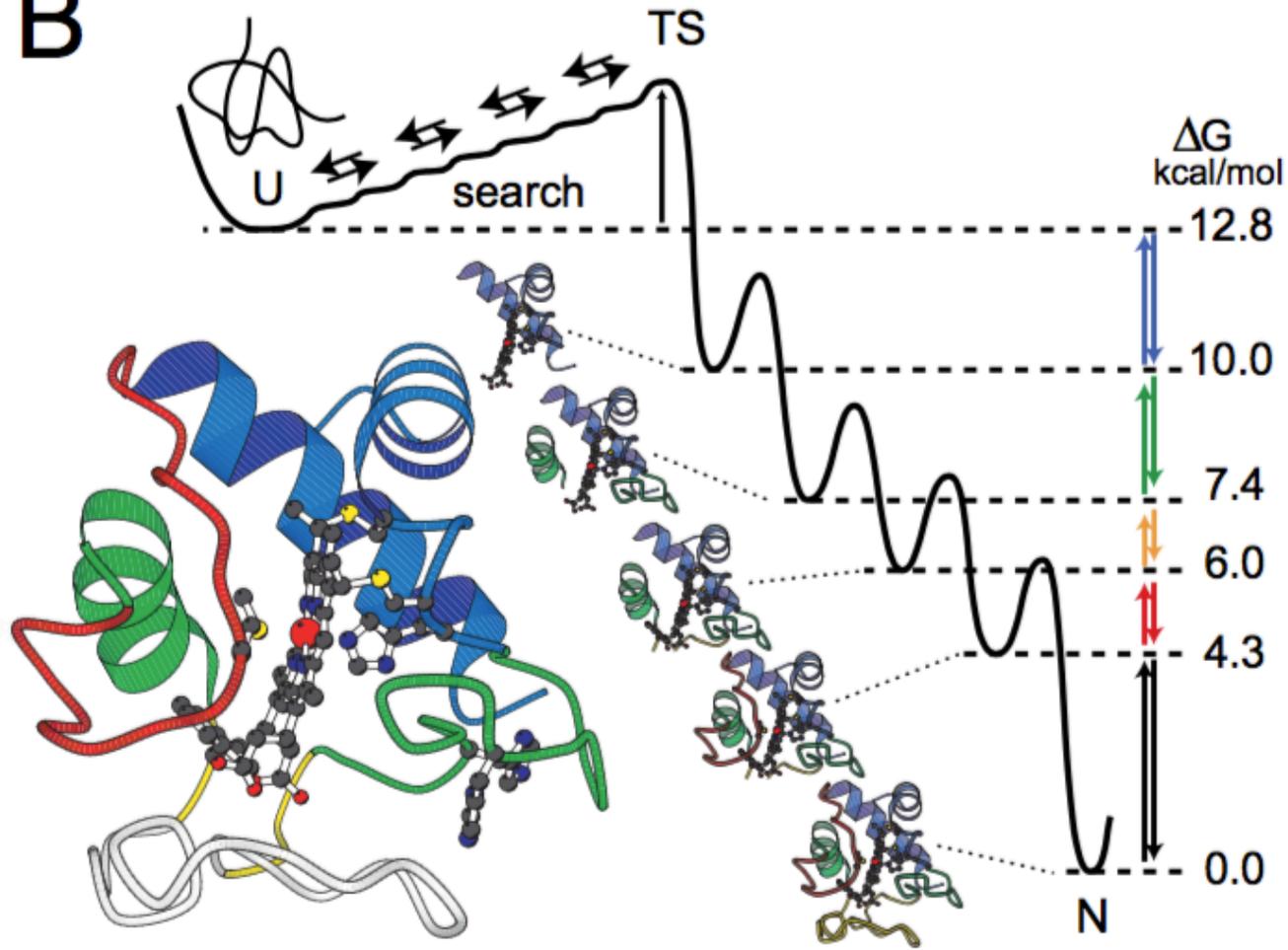
**Secondary Protein Structure**  
Local folding of the polypeptide chain into helices or sheets

**Tertiary Protein Structure**  
three-dimensional folding pattern of a protein due to side chain interactions

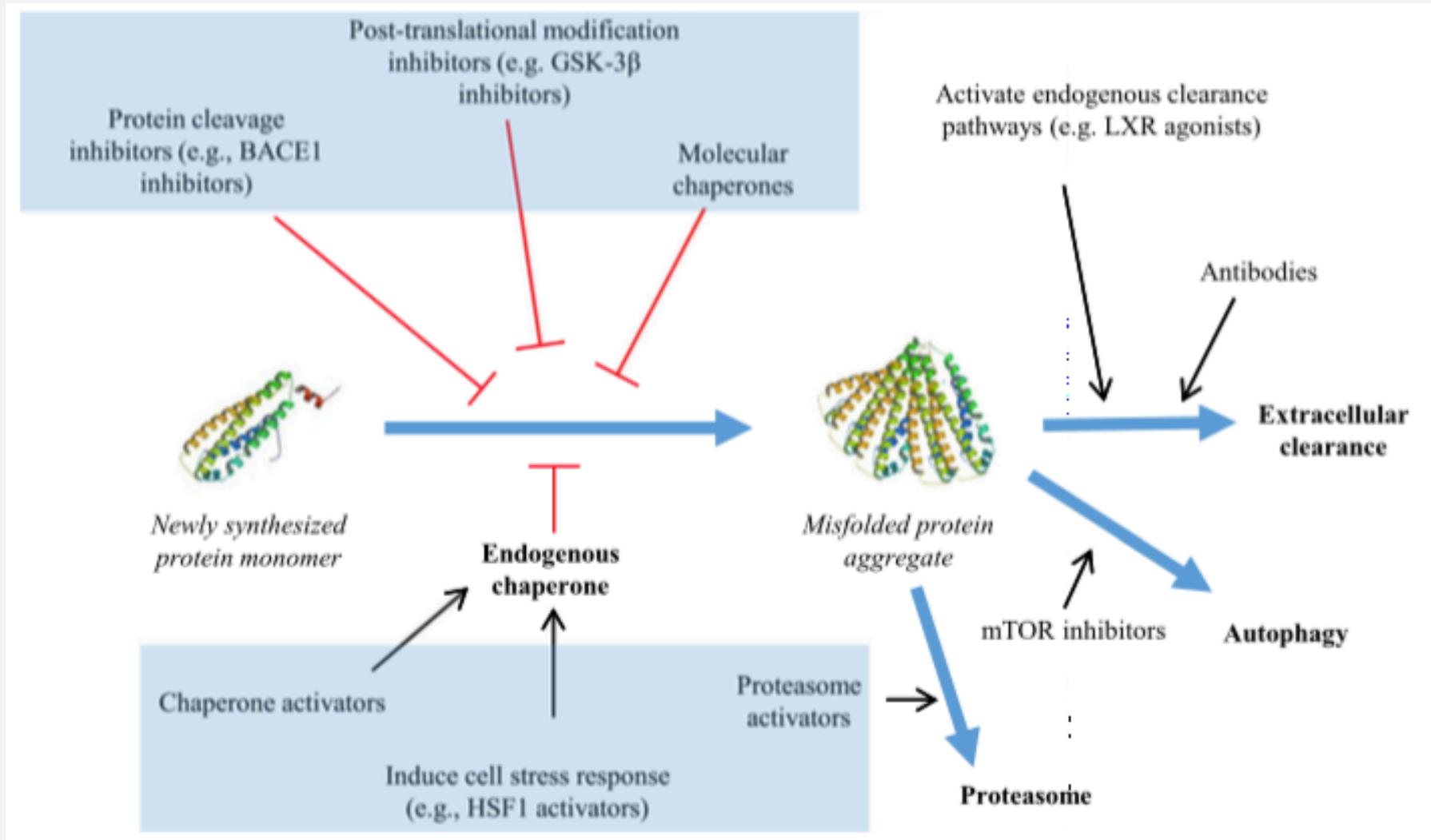
**Quaternary Protein Structure**  
protein consisting of more than one amino acid chain

agents: pH, temp, ionic strength, solubility



**A****B**

# Mechanisms involved in protein misfolding & therapeutic targets



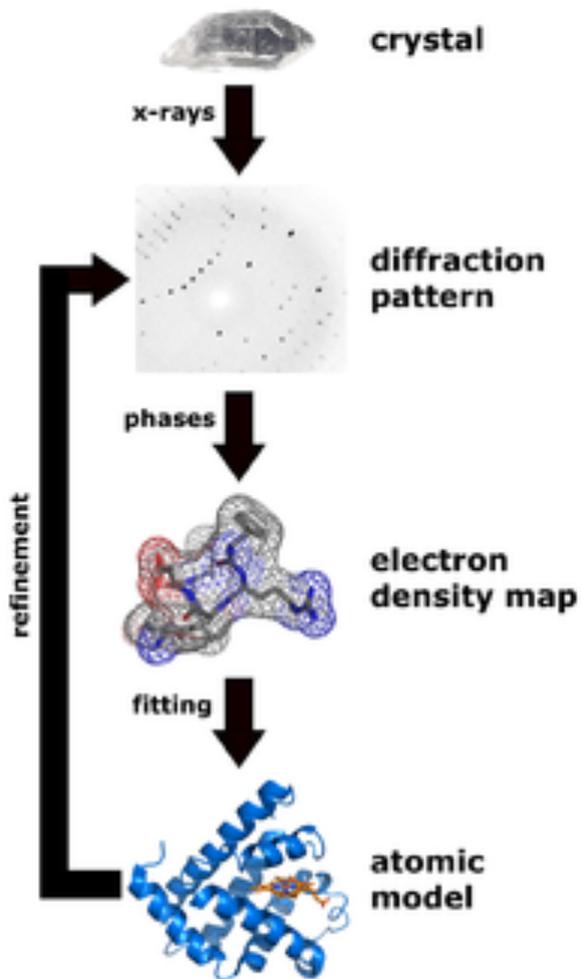
However, 50 y after C. B. Anfinsen showed that proteins can fold spontaneously without outside help, and despite the intensive work of thousands of researchers **leading to more than five publications per day in the current literature**, there is still no general agreement on the most primary questions.

**How do proteins fold?**

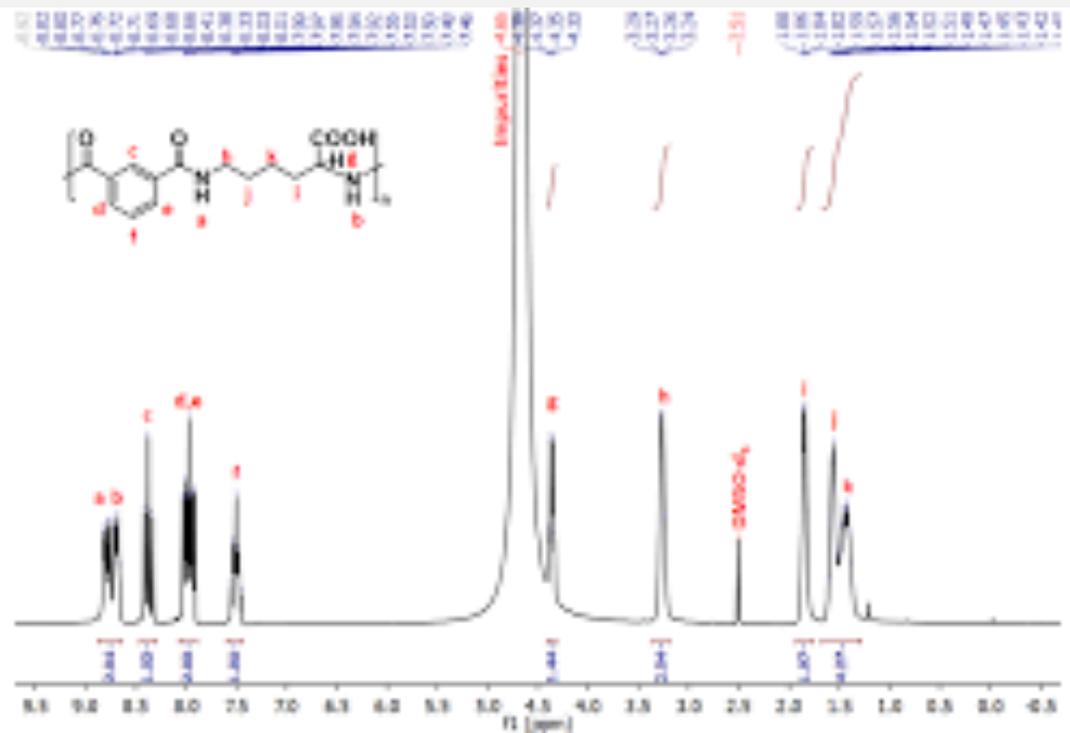
**Why do they fold in that way?**

**How is the course of folding encoded in a 1D amino acid sequence?**





X-ray  
crystallography



nuclear magnetic resonance  
spectroscopy

# COMPUTATIONAL METHODS FOR PROTEIN FOLD RECOGNITION

Computational methods for protein fold recognition can be generally categorized into three classes:

- (1) de novo modeling methods;
- (2) template-based methods; and
- (3) template-free methods.



Computational methods for protein fold recognition can be generally categorized into three classes:

- (1) **de novo modeling methods;** it requires long computational time and numerous sources,
- (2) **template-based methods; and**
- (3) **template-free methods.** and it can only be successfully applied in small proteins.



2. Template-based methods used to determine protein structures are based on the evolutionary relationships of proteins.



1. First, proteins of known structures retrieved from public protein structure databases (e.g., Protein Data Bank (PDB)) are used as template proteins for a query protein sequence.
2. Distant evolutionary relationships between a target sequence and proteins of known structure are detected.
3. To determine the optimal alignments, scoring functions are usually used as measures to evaluate the similarity between the profiles derived from a query protein and those of template proteins with known structures. *Z*-score and *E*-value are the two commonly used scoring functions.



4. 3D structure models based on template atom coordinates and optimal query-template alignments are built.
5. The optimal structure models are determined from the model candidates through further structure optimization. The commonly used structural optimization methods include energy minimization and loop modeling.

## Template-based methods

### 1. Fold and Function Assignment System (FFAS)

uses a profile-profile alignment strategy. Query and template profiles are obtained by PSI-BLAST searching against the NR85 database; these profiles are then aligned by a dot-product scoring function. The significance of alignment scores is calculated by comparing the protein with the distribution scores from pairs of unrelated protein

2. FFAS-3D, wherein is introduced structural information, such as secondary structure, solvent accessibility, and residue depth.



## template-based methods

1. FUGUE
2. Raptor
3. I-TASSER (Iterative Threading ASSEmbly Refinement),
4. ORION,
5. MODELLER
6. TMFR

Currently, CASP (Critical Assessment of protein Structure Prediction) is a mainstream platform used to establish an independent mechanism to assess the current methods employed in protein structure modeling.

( <http://predictioncenter.org/> )



## Template-based methods & problems

1. we need to determine the structures of template proteins. The three-dimensional structures of many proteins remain to be determined.
2. template-based modelling largely relies on the homology between target and template proteins. When the target and template proteins display a sequence similarity of  $>30\%$ , the use of sequence alignment methods (e.g., BLAST and SSEARCH) can reveal their evolutionary relationships. However, this approach is not available for non-obvious relationships between targets and templates with a sequence identity of lower than 20%-30%.
3. template-based structure modeling is time consuming. This approach always requires homology detection by searching target proteins against a template database to detect distant evolutionary relationships.



# Template-free methods

seek to build models and accurately predict protein structures solely based on amino acid sequences rather than on known structural proteins as templates.

1. Machine learning algorithms
2. Hidden Markov Model (HMM)
3. genetic algorithm
4. Artificial Neural Network
5. Support Vector Machines (SVMs), and
6. ensemble classifiers.

# Template-free methods

- the number of protein fold classes is limited.
- Machine learning aims to build a prediction model by learning the differences between different protein fold categories and use the learned model to automatically assign a query protein to a specific protein fold class.
- This approach is thus more efficient for large-scale predictions and can examine a large number of promising candidates for further experimental validation.

# Databases and Resources for Protein Folding

name	URL	description
<b>Protein Structure Databases</b>		
PDB	<a href="http://www.pdb.org">http://www.pdb.org</a>	macromolecular tertiary structure database
SCOP	<a href="http://scop.mrc-lmb.cam.ac.uk/scop">http://scop.mrc-lmb.cam.ac.uk/scop</a>	structure classification of proteins
CATH	<a href="http://www.cathdb.info">http://www.cathdb.info</a>	HMM and domain protein 3D structure classification
<b>Thermodynamic and Kinetic Databases</b>		
ProTherm	<a href="http://www.abren.net/protherm/">http://www.abren.net/protherm/</a>	protein thermodynamic database
KineticDB	<a href="http://kineticdb.protres.ru/db/index.pl">http://kineticdb.protres.ru/db/index.pl</a>	manually curated databases of kinetic data
Muñoz lab	<a href="http://tmg.cib.csic.es/servers/data-tables">http://tmg.cib.csic.es/servers/data-tables</a>	database of kinetic rates for proteins and their mutants
PFD	<a href="http://pfd.med.monash.edu">http://pfd.med.monash.edu</a>	kinetic database collecting folding rates and energies
REFOLD	<a href="http://refold.med.monash.edu.au">http://refold.med.monash.edu.au</a>	database of optimized refolding protocols

Biochemistry 2013, 52, 8601–8624



name	URL	description
<b>Protein Structure Prediction Tools and Resources</b>		
I-TASSER	<a href="http://zhanglab.ccmb.med.umich.edu/I-TASSER">http://zhanglab.ccmb.med.umich.edu/I-TASSER</a>	structure prediction by threading
ModBase	<a href="https://modbase.compbio.ucsf.edu/scgi/modweb.cgi">https://modbase.compbio.ucsf.edu/scgi/modweb.cgi</a>	repository of models predicted by homology
MODELER	<a href="http://www.salilab.org/modeller">http://www.salilab.org/modeller</a>	standard homology modeling tool
Protein Model Portal	<a href="http://www.proteinmodelportal.org">http://www.proteinmodelportal.org</a>	resources and services for protein structure prediction
ROBETTA	<a href="http://rosetta.bakerlab.org">http://rosetta.bakerlab.org</a>	<i>de novo</i> and homology modeling algorithm
<b>Physics-Based Energy Functions</b>		
AMBER	<a href="http://amber.scripps.edu">http://amber.scripps.edu</a>	molecular mechanics force field and package for simulations
CHARMM	<a href="http://www.charmm.org">http://www.charmm.org</a>	empirical atomic force fields for molecular dynamics
GROMOS	<a href="http://www.igc.ethz.ch/gromos">http://www.igc.ethz.ch/gromos</a>	energy function included in GROMACS
<b>Knowledge-Based Potentials</b>		
ANOLEA	<a href="http://protein.bio.puc.cl/cardex/servers/anolea">http://protein.bio.puc.cl/cardex/servers/anolea</a>	atomic statistical potential scoring nonlocal interactions
DFIRE	<a href="http://sparks.informatics.iupui.edu/yueyang/DFIRE">http://sparks.informatics.iupui.edu/yueyang/DFIRE</a>	residue-specific and distance-scaled mean force potential
PROSA-web	<a href="https://prosa.services.came.sbg.ac.at">https://prosa.services.came.sbg.ac.at</a>	knowledge-based potential for scoring protein structures
<b>Prediction of Protein Stability</b>		
AUTO-MUTE	<a href="http://proteins.gmu.edu">http://proteins.gmu.edu</a>	machine learning and statistical potential for $\Delta\Delta G$ predictions
CUPSAT	<a href="http://cupsat.tu-bs.de">http://cupsat.tu-bs.de</a>	statistical potentials for the prediction of $\Delta\Delta G$
DMUTANT	<a href="http://sparks.informatics.iupui.edu/hzhou/mutation.html">http://sparks.informatics.iupui.edu/hzhou/mutation.html</a>	prediction of $\Delta\Delta G$ using DFIRE statistical potential
Fold-X	<a href="http://foldx.crg.es">http://foldx.crg.es</a>	empirical scoring function for the prediction of protein stability
I-Mutant	<a href="http://folding.biofold.org/i-mutant">http://folding.biofold.org/i-mutant</a>	sequence and structure SVM-based method
PopMusic	<a href="http://babylone.ulb.ac.be/popmusic">http://babylone.ulb.ac.be/popmusic</a>	neural network and statistical potential for $\Delta\Delta G$ predictions
PreThermut	<a href="http://www.mobioinfor.cn/prethermut">http://www.mobioinfor.cn/prethermut</a>	random forest for single- and multiple-mutation predictions
ProMaya	<a href="http://bental.tau.ac.il/ProMaya">http://bental.tau.ac.il/ProMaya</a>	random forest and filtering model for $\Delta\Delta G$ predictions
MuPro	<a href="http://mupro.proteomics.ics.uci.edu">http://mupro.proteomics.ics.uci.edu</a>	structure-based SVM for $\Delta\Delta G$ predictions

# Methods and Tools for Protein Folding

Biochemistry 2013, 52, 8601–8624



# Tools for Protein Aggregation Available on the Web

name	URL	description
Empirical Methods		
AGGRESKAN	<a href="http://bioinf.uab.es/aggrescan">http://bioinf.uab.es/aggrescan</a>	aggregation propensity scale from <i>in vivo</i> experiments
Tango	<a href="http://tango.crg.es">http://tango.crg.es</a>	physicochemical properties of $\beta$ -sheet formation in core regions
Zyggregator	<a href="http://www-vendruscolo.ch.cam.ac.uk/zyggregator.php">http://www-vendruscolo.ch.cam.ac.uk/zyggregator.php</a>	physicochemical propensities of residues
Structure-Based Tools		
BETASCAN	<a href="http://groups.csail.mit.edu/cb/betascan">http://groups.csail.mit.edu/cb/betascan</a>	$\beta$ -strands and strand pair scores from parallel $\beta$ -sheet
FoldAmyloid	<a href="http://bioinfo.protres.ru/fold-amyloid/oga.cgi">http://bioinfo.protres.ru/fold-amyloid/oga.cgi</a>	hydrogen bond probability and residue packing density
Net-CSSP	<a href="http://cssp2.sookmyung.ac.kr">http://cssp2.sookmyung.ac.kr</a>	$\beta$ -strand propensity from buried and highly interacting regions
PASTA	<a href="http://protein.cribi.unipd.it/pasta">http://protein.cribi.unipd.it/pasta</a>	$\beta$ -parallel and antiparallel scores for amyloid formations
Waltz	<a href="http://waltz.switchlab.org">http://waltz.switchlab.org</a>	position-specific scoring matrix to predict aggregating regions

Biochemistry 2013, 52, 8601–8624



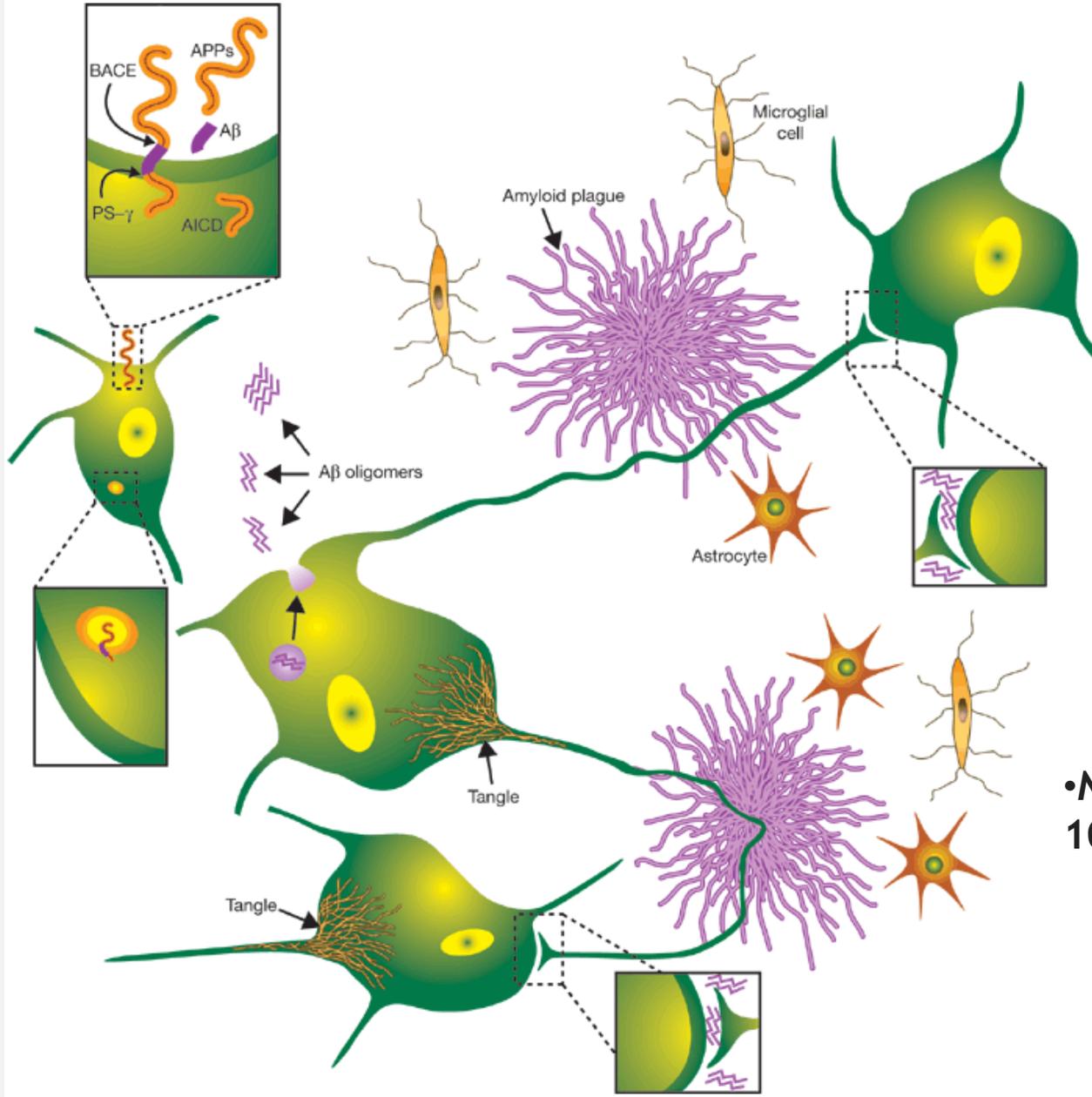
**WORK IN PROGRESS**

**IN SILICO STUDY OF THE  
TERTIARY STRUCTURE OF WELL-  
ESTABLISHED MUTATIONS IN  
ALZHEIMER'S DISEASE**

# Alzheimer's Disease is a folding disorder

- Irreversible progressive dementia with long prodromal stages and up to now there is a lack of effective pharmacotherapy options
- evaluation of the prediction algorithms
- While the clinical symptoms of the disease are defined by cognitive impairment, the causes leading to memory decline are strongly tied to deposits of misfolded protein aggregates.





•*Nature Cell Biology* 6, 1054-1061 (2004)

# *In Silico* Analysis of the Apolipoprotein E and the Amyloid $\beta$ Peptide Interaction: Misfolding Induced by Frustration of the Salt Bridge Network

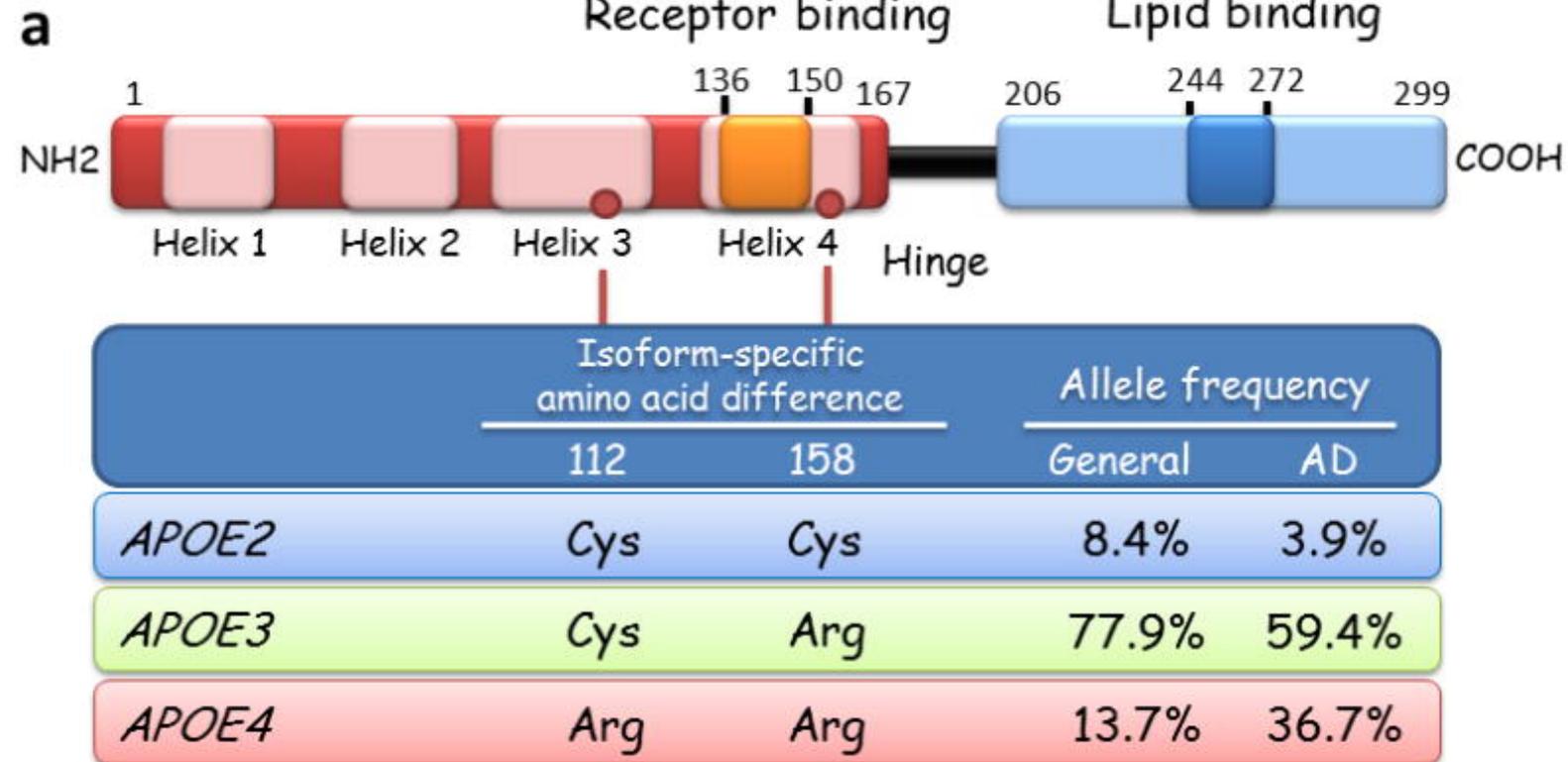
Jinghui Luo<sup>1</sup>, Jean-Didier Maréchal<sup>2</sup>, Sebastian Wärmländer<sup>1</sup>, Astrid Gräslund<sup>1</sup>, Alex Perálvarez-Marín<sup>1,3\*</sup>✉

**1** Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden, **2** Unitat de Química Física, Departament de Química, Universitat Autònoma de Barcelona, Bellaterra, Spain, **3** Unitat de Biofísica, Departament de Bioquímica i de Biologia Molecular i Centre d'Estudis en Biofísica, Universitat Autònoma de Barcelona, Bellaterra, Spain

## Abstract

The relationship between Apolipoprotein E (ApoE) and the aggregation processes of the amyloid  $\beta$  ( $A\beta$ ) peptide has been shown to be crucial for Alzheimer's disease (AD). The presence of the ApoE4 isoform is considered to be a contributing risk factor for AD. However, the detailed molecular properties of ApoE4 interacting with the  $A\beta$  peptide are unknown, although various mechanisms have been proposed to explain the physiological and pathological role of this relationship. Here, computer simulations have been used to investigate the process of  $A\beta$  interaction with the N-terminal domain of the human ApoE isoforms (ApoE2, ApoE3 and ApoE4). Molecular docking combined with molecular dynamics simulations have been undertaken to determine the  $A\beta$  peptide binding sites and the relative stability of binding to each of the ApoE isoforms. Our results show that from the several ApoE isoforms investigated, only ApoE4 presents a misfolded intermediate when bound to  $A\beta$ . Moreover, the initial  $\alpha$ -helix used as the  $A\beta$  peptide model structure also becomes unstructured due to the interaction with ApoE4. These structural changes appear to be related to a rearrangement of the salt bridge network in ApoE4, for which we propose a model. It seems plausible that ApoE4 in its partially unfolded state is incapable of performing the clearance of  $A\beta$ , thereby promoting amyloid forming processes. Hence, the proposed model can be used to identify potential drug binding sites in the ApoE4- $A\beta$  complex, where the interaction between the two molecules can be inhibited.

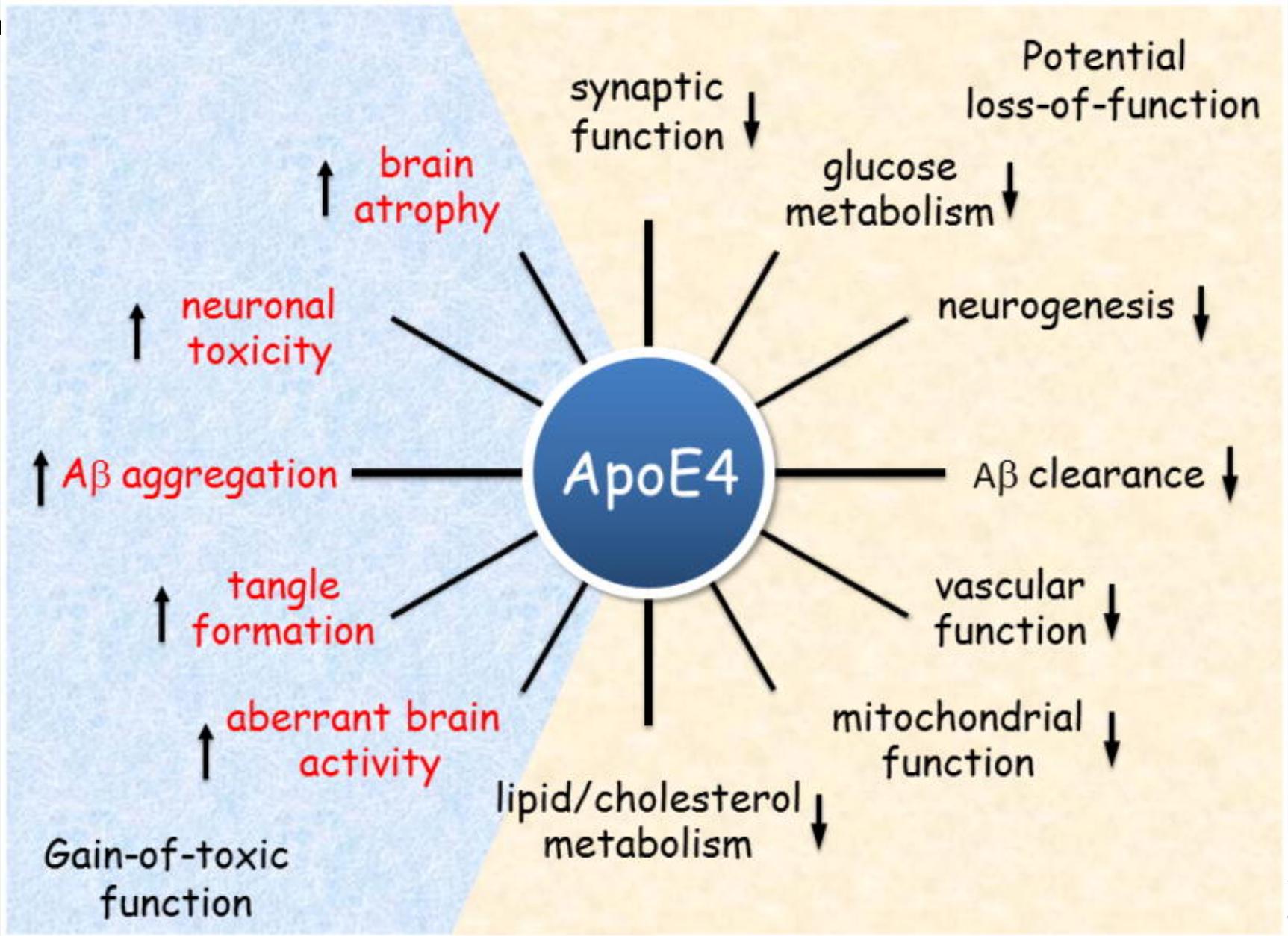




**b**

	<i>APOE4</i>		
	Non-carrier	Heterozygous	Homozygous
AD frequency	20%	47%	91%
Mean age of clinical onset	84-yr	76-yr	68-yr

Nat Rev Neurol. 2013  
Feb; 9(2): 106–118.



Nat Rev  
Neurol.  
2013 Feb;  
9(2): 106–  
118.

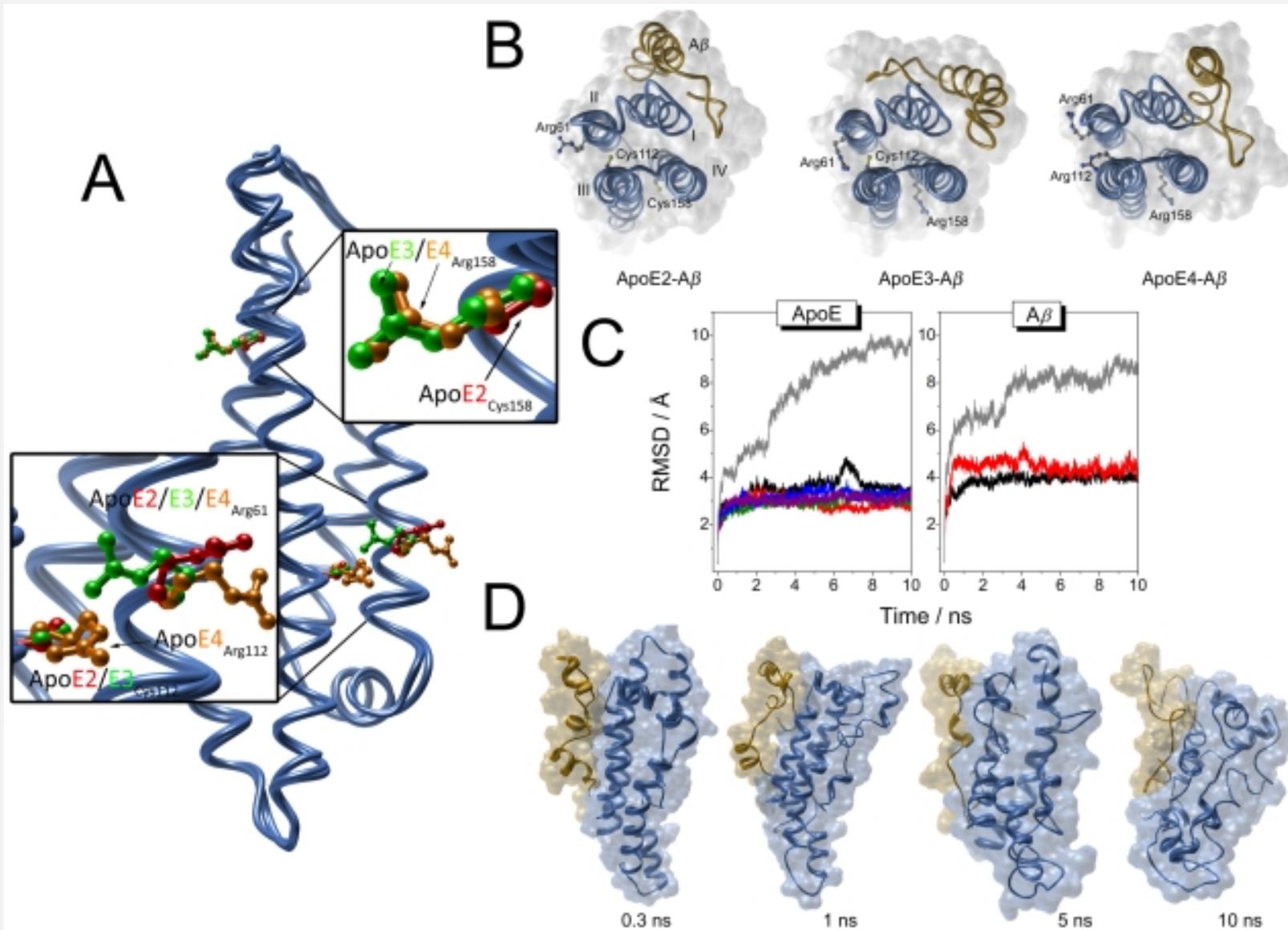


Computer simulations have been used to investigate the process of  $A\beta$  interaction with the N-terminal domain of the human ApoE isoforms (ApoE2, ApoE3 and ApoE4). Molecular docking combined with molecular dynamics simulations have been undertaken to determine the  $A\beta$  peptide binding sites and the relative stability of binding to each of the ApoE isoforms.

[PLoS Comput Biol.](#) 2010 Feb; 6(2): e1000663.



# Apolipoprotein E (ApoE) and the aggregation processes of the amyloid $\beta$ ( $A\beta$ ) peptide



The model presented has implications for therapeutic drug design for AD, as it defines on a molecular level the ApoE- $A\beta$  complex as a potential drug target

[PLoS Comput Biol.](#) 2010 Feb; 6(2): e1000663.



# Hypothesis

**What is the connection between the different isoforms of well known proteins implicated in the pathogenesis and progression of Alzheimer's Disease?**

**Could we find a correlation between their misfolded forms and the events occurring at the molecular level?**

# Aim

- Structural algorithm evaluation on experimentally determined structures
- Algorithm implementation to unknown structures
- Structural alignment between the proteins
- Classify protein tertiary structures based on their similarity
- Underline the mutation footprint on the 3d structures

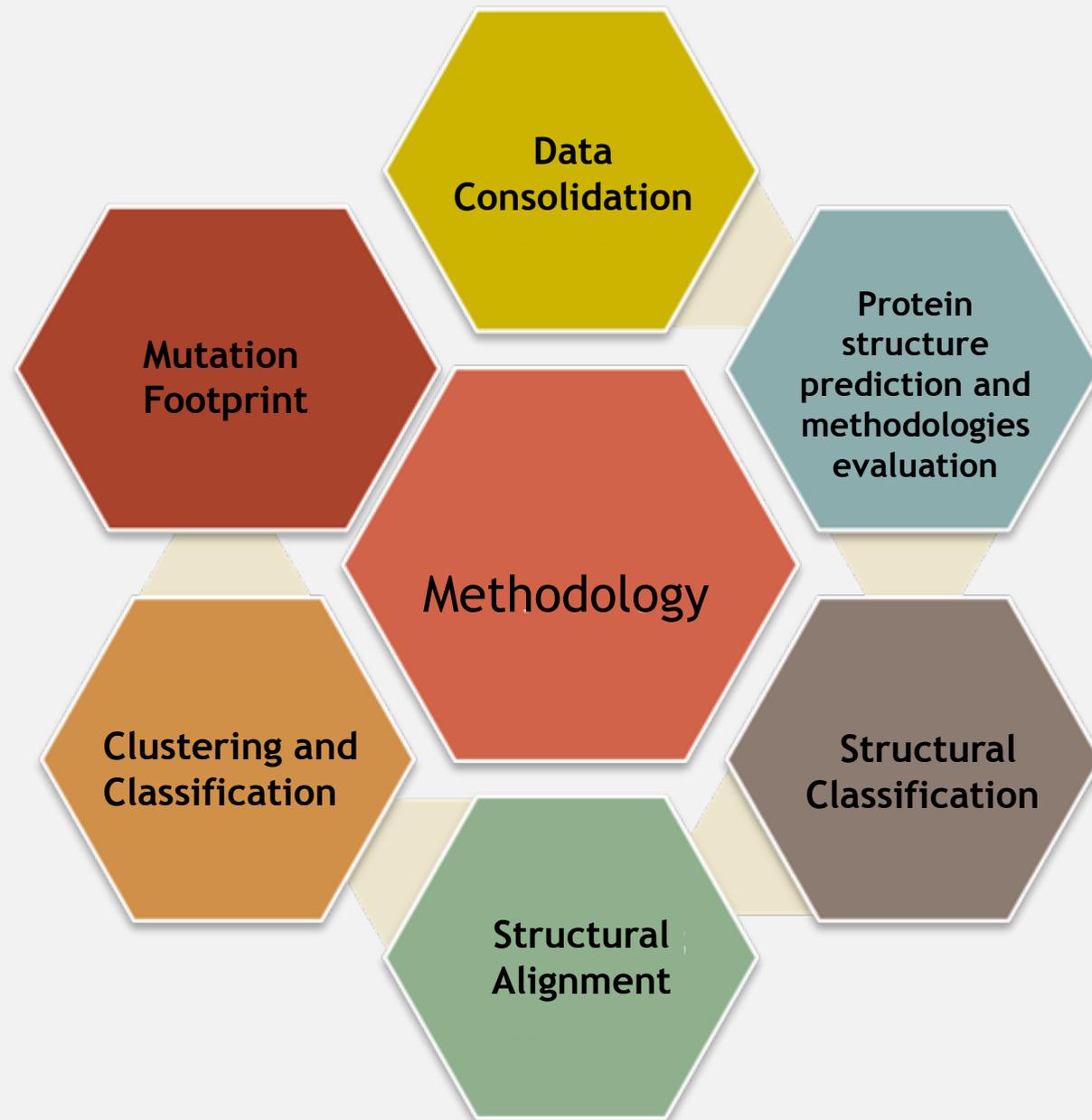
# Inovation

- ability to highlight broad ensembles of genes that drive key disease mechanisms
- These genes may hold strong potential for the identification of promising therapeutic targets.

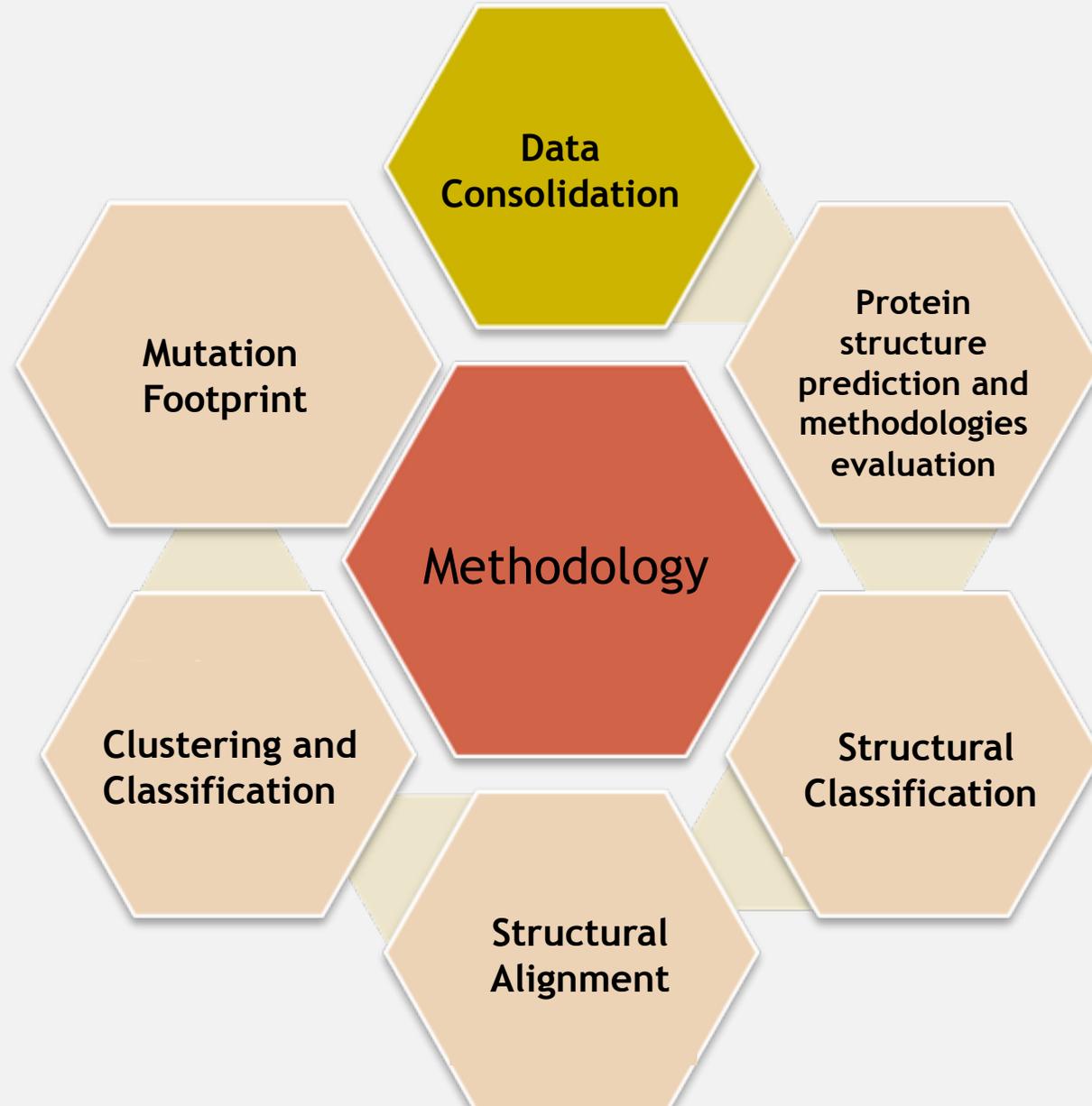
## Methodology in summary

- study of the experimentally determined protein structures,
- evaluation of the prediction algorithms
- prediction of the AD related proteins with unknown structures
- evaluation of the mutation footprint of the established mutations in the tertiary structure.

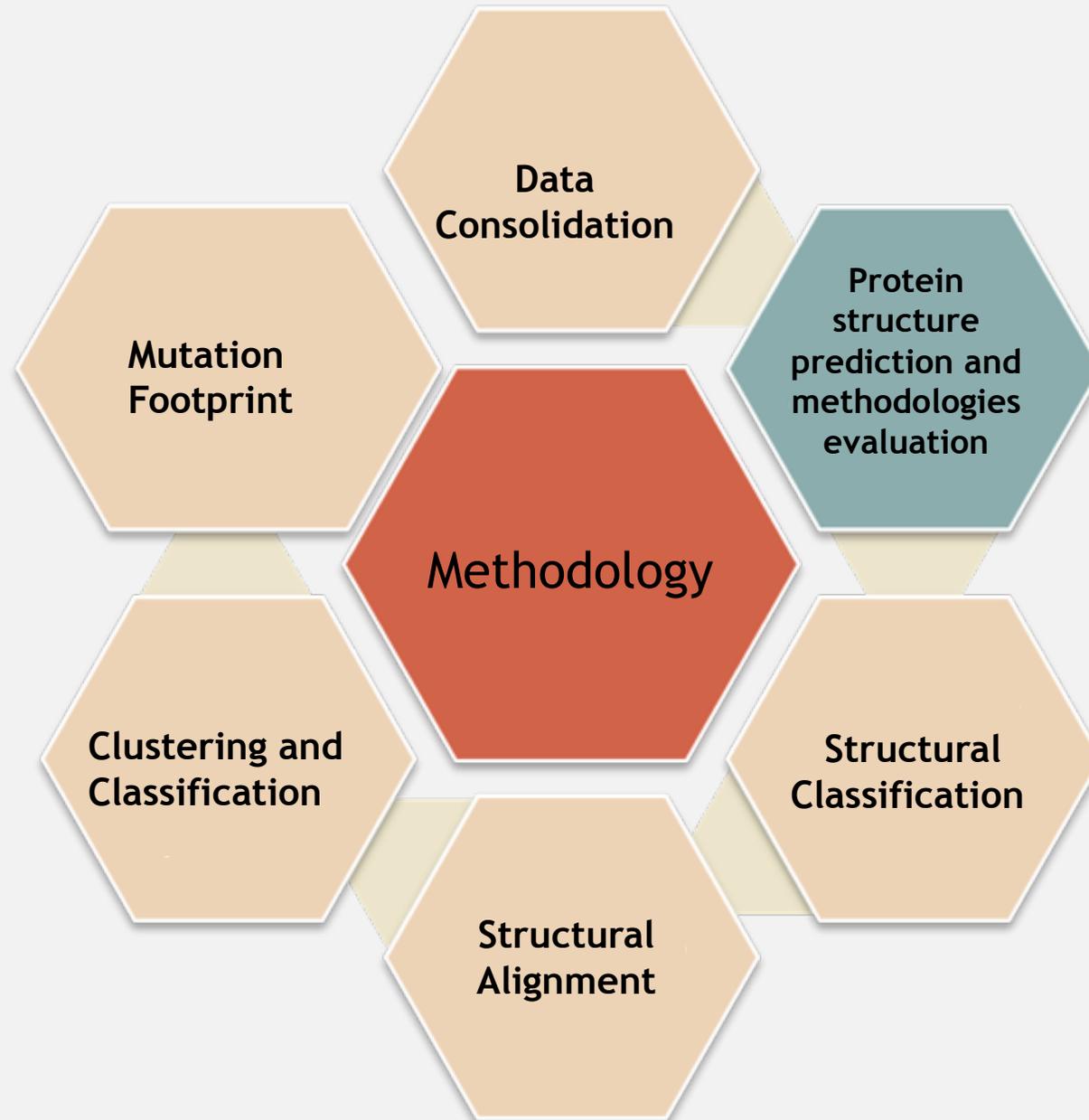
# Methodology



# Methodology



# Methodology



The evaluation method is included in the following steps:

1. Experimentally-determined structures of proteins from PDB will be utilized as guides.
2. Prediction of the tertiary structure for all the guide-proteins, by deploying only the primary structure.
3. Similarity metrics are then applied in the predicted structures in order to compare the predicted models with the experimentally-determined.

The similarity metrics that are **utilized are RMSD and TM-score**. Consequently, methods beyond classic bioinformatics (for example structure similarity following computer-vision based algorithms) is utilized for more detailed feature extraction.



# The evaluation method is included in the following steps:

1. Experimentally-determined structures of proteins from PDB will be utilized as guides.
2. Prediction of the tertiary structure for all the guide-proteins, by deploying only the primary structure.
3. Similarity metrics are then applied in the predicted structures in order to compare the predicted models with the experimentally-determined.

The similarity metrics that are **utilized** are **RMSD** and **TM-score**. Consequently, methods beyond classic bioinformatics (for example structure similarity following computer-vision based algorithms) is utilized for more detailed feature extraction.



Software	Method	Description	Category
<b>MODELLER</b>	Satisfaction of spatial restraints	Standalone program mainly in Fortran and Python	homology
<b>SWISS-MODEL</b>	Local similarity/fragment assembly	Automated webserver (based on ProModII)	homology
<b>HHpred</b>	Template detection, alignment, 3D modeling	Interactive webserver with help facility	threading
<b>I-TASSER</b>	Combination of ab initio folding and threading methods	Structural and function predictions	Combination
<b>ROBETTA</b>	Rosetta homology modeling and ab initio fragment assembly with GinzU domain prediction	Webserver	Ab initio
<b>QUARK</b>	Monte Carlo fragment assembly	On-line server for protein modeling (best for ab initio folding in CASP9)	Ab initio

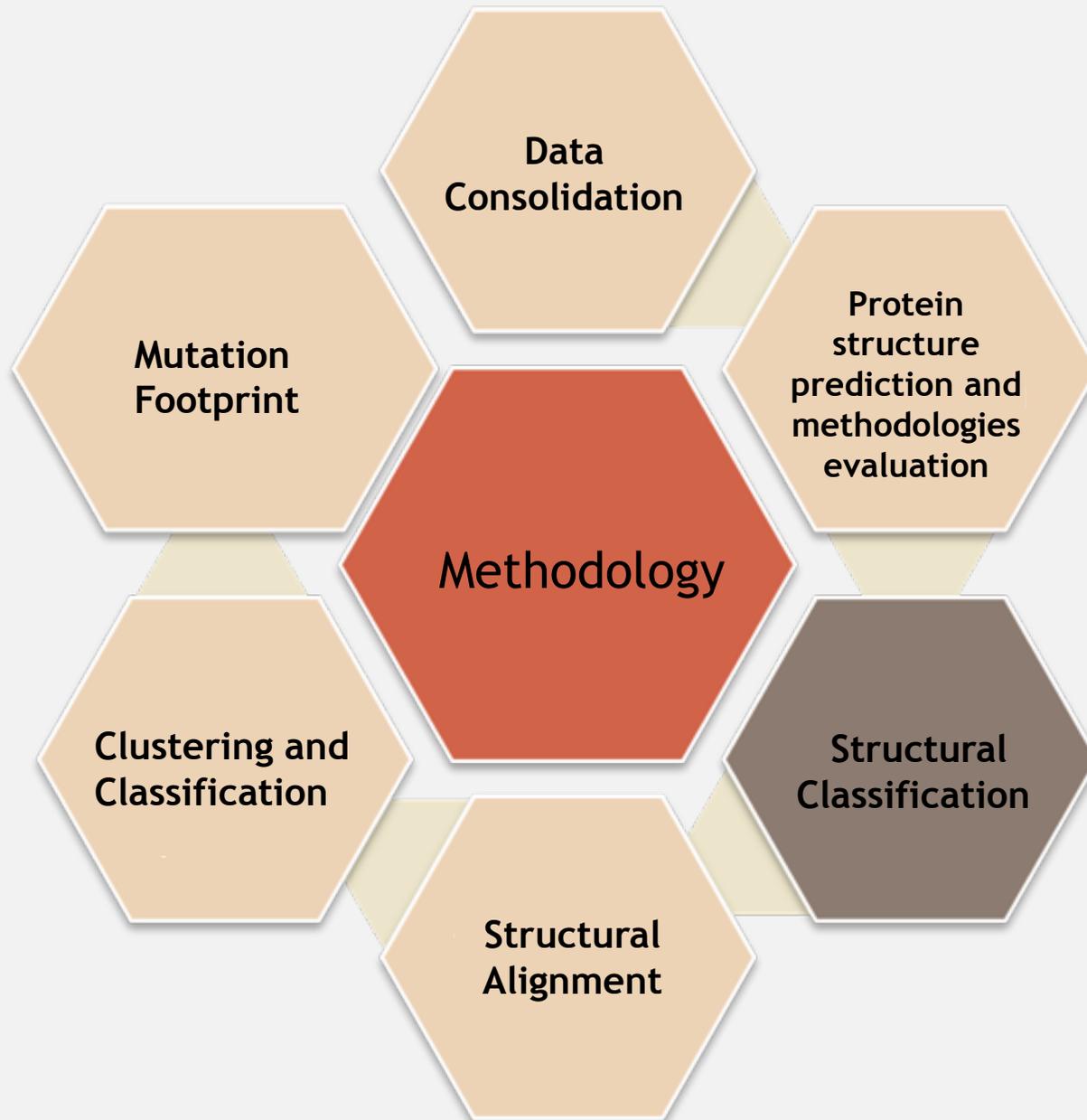
# The evaluation method is included in the following step:

1. Experimentally-determined structures of proteins from PDB will be utilized as guides.
2. Prediction of the tertiary structure for all the guide-proteins, by deploying only the primary structure.
3. Similarity metrics are then applied in the predicted structures in order to compare the predicted models with the experimentally-determined.

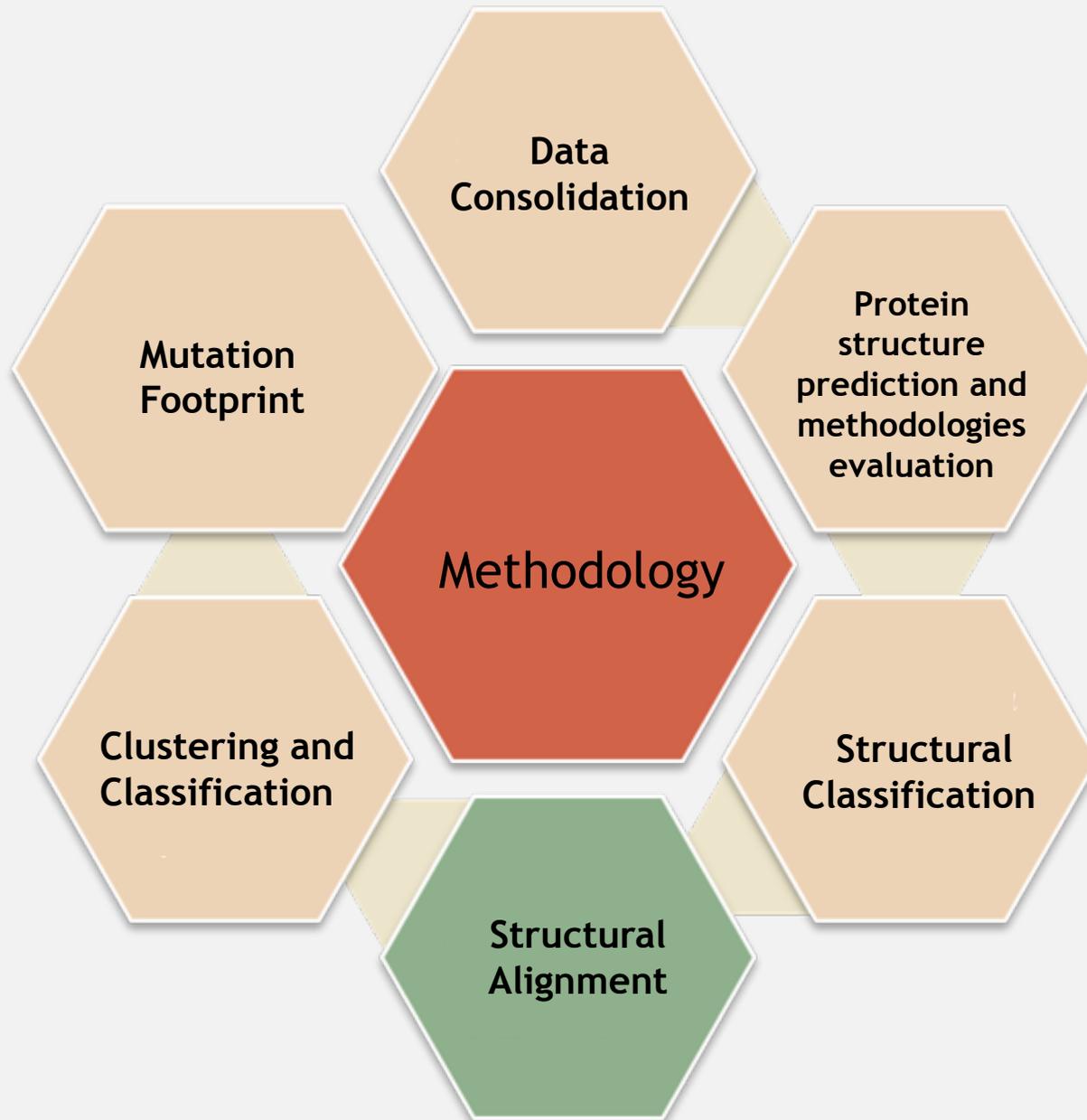
The similarity metrics that are **utilized are RMSD and TM-score**. Consequently, methods beyond classic bioinformatics (for example structure similarity following computer-vision based algorithms) is utilized for more detailed feature extraction.



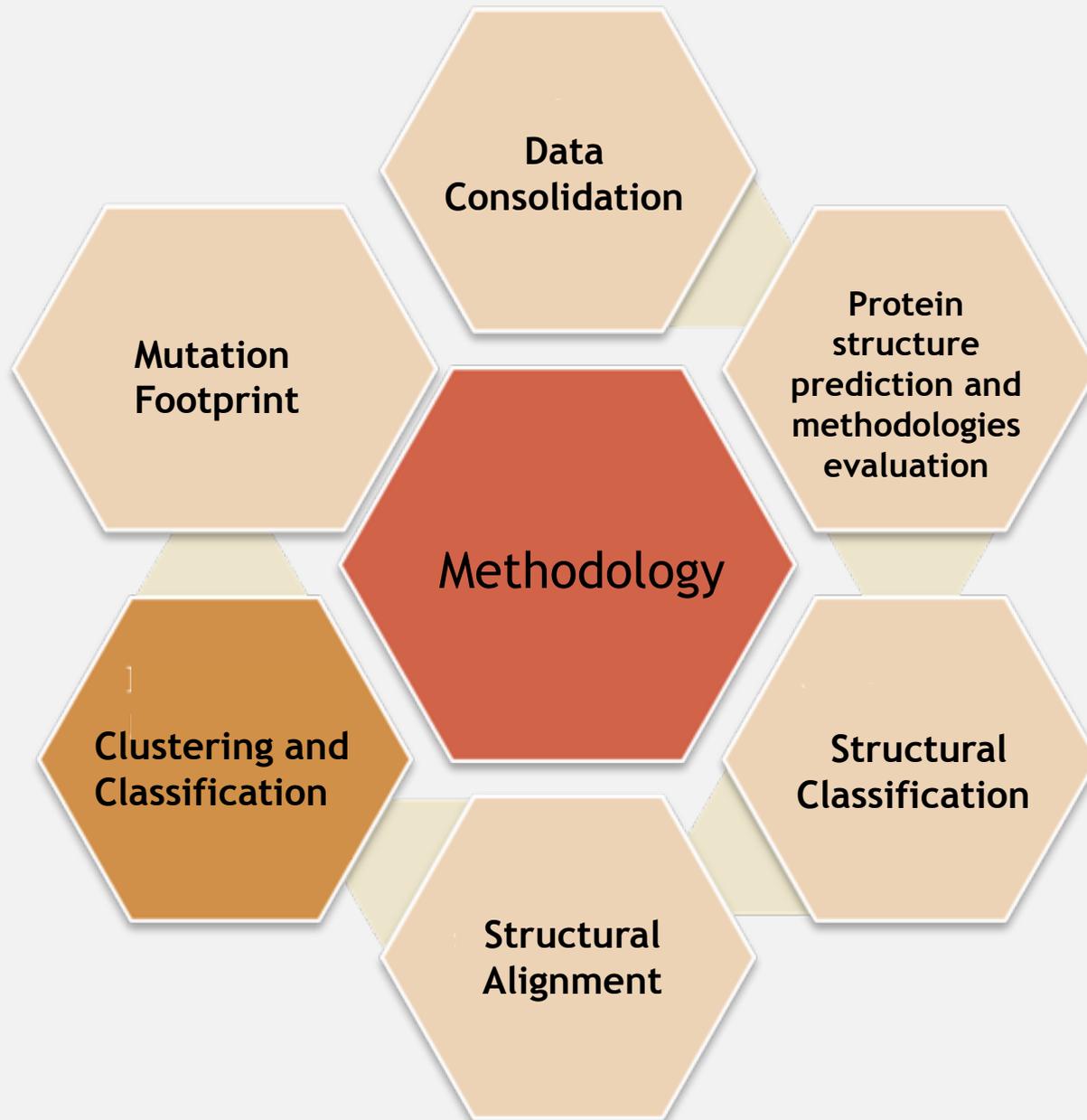
# Methodology



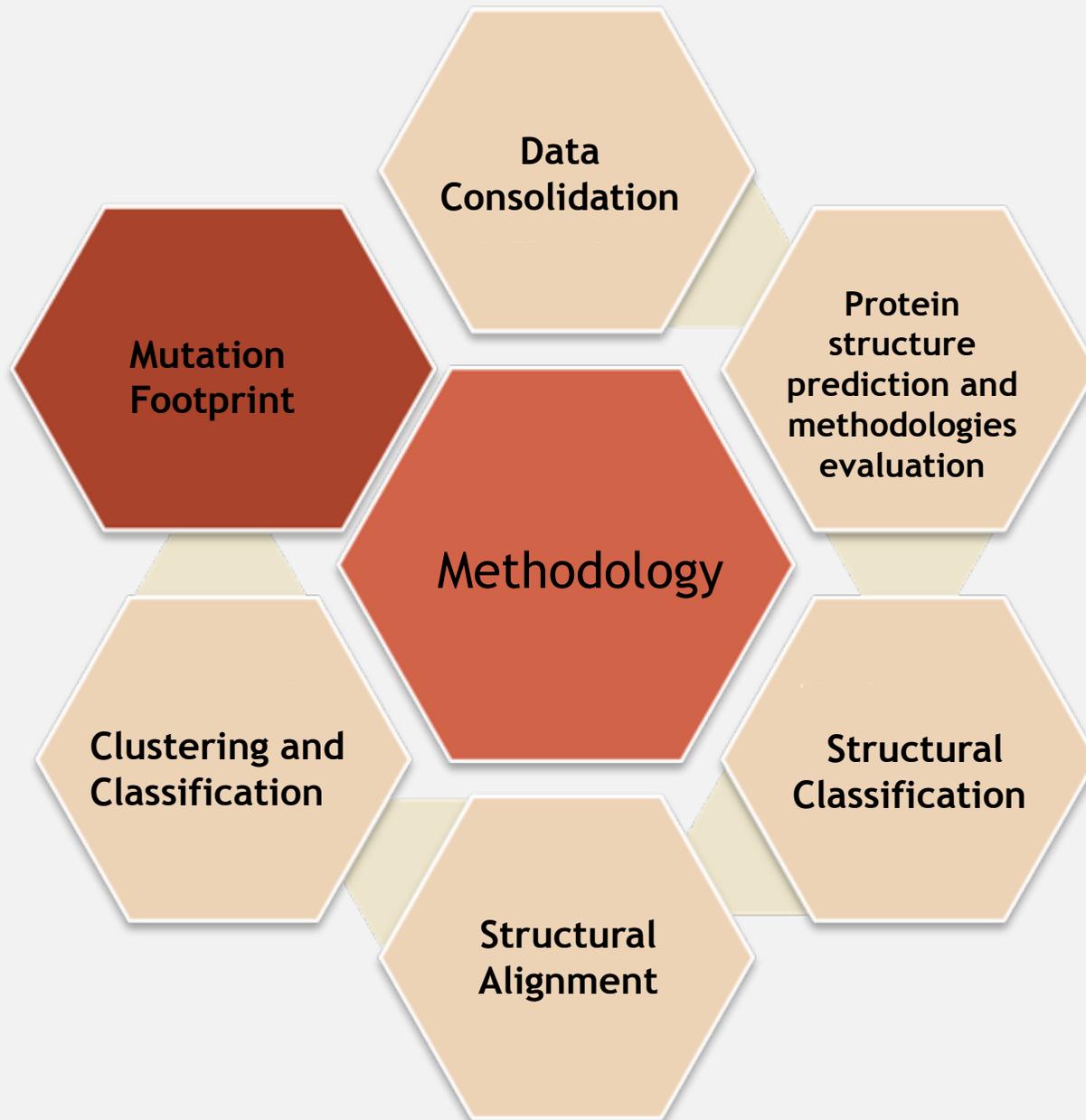
# Methodology



# Methodology



# Methodology



# Summary

- Risk prediction standpoint,  
networks exhibiting coexistent genetic variation and biological perturbation would represent prime targets in the development of personalized, burden-based genetic susceptibility tests.
- Therapeutic strategies development  
pathways and networks displaying multi-omics relationships in AD would reduce the search space for rational drug design and may highlight “hub” genes for therapeutic cocktail approaches, such as in the polypharmacy strategies successfully employed for AIDS and various cancers.



Foldit is a revolutionary crowdsourcing computer game enabling *you* to contribute to important scientific research.



**Protein structure prediction**

**Protein design**

<https://fold.it/portal/>

