In Silico Search for Endogenous Inhibitors of Protein Misfolding

Weaver, D.F.[1,2,3], Meek, A.[1], Gupta, M.[1], Wang, Y.F.[1]

[1]Krembil Research Institute, University Health Network, Toronto, Canada

[2]Department of Chemistry, University of Toronto, Canada

[3]Department of Pharmaceutical Sciences, University of Toronto, Canada

1. Introduction

Protein misfolding is a fundamental disease process implicated in many human disorders, particularly dementias such as Alzheimer's disease, but also in diabetes and specific types of heart and kidney failure. Proteins are the structural and functional workhouse molecules of the human body – beneficial activities that are dependent upon the protein being folded into a correct shape. Since protein shape is central to health, it is reasonable to postulate the existence of compounds endogenous to the human body which ameliorate the pathological consequences of protein misfolding; the concept of searching for endogenous anti-protein misfolding compounds is unique. We have used an *in silico* high throughput screen of small molecules endogenous in the human brain to identify multiple classes of agents capable of inhibiting the aberrant protein misfolding implicated in the pathogenesis of Alzheimer's dementia. We then extend this discovery to show that these endogenous compounds also inhibit the misfolding of proteins implicated in diabetes, thereby demonstrating that these agents are not disease specific and are applicable to multiple classes of protein misfolding diseases, including two of the most significant disorders afflicting humankind, namely dementia and diabetes.

2. Background

Protein misfolding (*i.e.* proteopathy) is a poorly understood, yet fundamental, pathological mechanism contributing to a wide variety of common and chronic human diseases. Cellular death is an ultimate consequence of most disease mechanisms including proteopathy. The final common pathway for cell death involves calcium dysregulation arising from either necrotic or apoptotic processes. In turn, necrosis and apoptosis are a product of one of more pathologies including developmental errors, degenerative disorders, ischaemia, infection/inflammation, neoplasia, nutritional/ metabolic disorders, chemical toxicity or trauma (*i.e.* "DINT[2]" processes). To this comprehensive array of possible pathological processes we add protein misfolding.

Proteins are central to the architecture and functions of the human body; accordingly, proteins are quantitatively the primary macromolecule. Some proteins serve a structural role (*e.g.* collagen) others catalyse biochemical processes (*e.g.* enzymes). The capacity of proteins to fulfil these diverse roles is directly dependent upon overall protein conformation, which itself is a product of a correct folding process

whereby the primary amino acid sequence is converted into a structurally sound shape. If an error occurs during the folding process, the resulting misfolding leads to dysfunctional proteins and thus disease.

Many diseases arise in whole or in part from protein misfolding. The diseases most commonly associated with protein misfolding include Alzheimer's disease (misfolding of β-amyloid and tau)[1,2], Parkinson's disease (α-synuclein), frontotemporal dementia (tau), chronic traumatic encephalopathy (tau), and progressive supranuclear palsy (tau). However, proteopathy is not restricted to brain-based diseases and contributes to multiple other non-neurological disorders, a partial listing of which includes: type II diabetes (amylin), sickle cell anaemia (haemoglobin), cardiac amyloidosis (atrial naturetic factor), dialysis amyloidosis (beta2-microglobulin) and cataracts (crystallins). Of these disorders, Alzheimer's disease (AD) is emerging as the most significant proteopathy afflicting millions of people worldwide.

3. Hypothesis and Aim

It is reasonable to hypothesize the existence of compounds endogenous to the human body that could disrupt or interrupt the pathogenesis of proteopathies. Since protein function is dependent on a three-dimensional structure achieved via folding on physiological timescales, cells have evolved internal surveillance mechanisms to maintain normal protein conformation and to monitor for misfolding. Thus, it is possible that a proteopathic disease such as AD may have its natural history modulated by one or more endogenous molecules. The aim of this study is to use *in silico* screening of a library of endogenous brain molecules to identify a compound capable of inhibiting the misfolding of Aβ and/or tau.

4. Methods

Searching for an endogenous anti-AD compound requires: 1. a comprehensive library of molecules naturally occurring in the human brain, and 2. a computational model of a receptor capable of inhibiting Aβ-mediated proteopathy.

We initially assembled a library of 1,100 human brain molecules (molecular weight < 650 g/mol). Next, the model receptor was identified. Giulian *et al.* (1998)[3] proposed that the $H_{13}HQK_{16}$ motif with Aβ provided a structural basis for the proteopathy of AD being essential for Aβ aggregation via a heparin sulfate proteoglycan mediated cell surface mechanism. We noticed and analogous BBXB amino acid motif, where "B" is a basic amino acid and X is any other amino acid, in tau ($K_{141}KAK_{144}$).[4] The endogenous brain compounds were thus screened against the Aβ's $E_{11}VHHQKLV_{18}$ domain using a molecular dynamics approach. Molecular modelling optimizations were performed using the AMBER10:EHT force field in MOE versions 2016.08 and 2016.08.02.[5] Solvation was simulated using the reaction field model. Since a crystal

structure of Aβ is not available, NMR structures available in the RCSB protein data bank were used. From the available structures, the 1IYT conformation of Aβ was selected: it is composed of 42 amino acid residues with two α-helices separated by a sharp hydrogen bonded turn, and its length makes it more prone to aggregation. The Aβ model was corrected for any structural anomalies, and the histidine residues were charged to reflect the microenvironment that would surround the protein in the brain. Each of the identified endogenous brain compounds was built and geometry optimized to an energy minimum. The charged functional or aromatic groups were oriented 3.0 Å from any two of the charged or aromatic amino acid side-chains in the EVHHQKLV region of Aβ for the 1IYT conformer. Each system was then optimized to an energy minimum. The final orientation of the ligand relative to the protein was recorded along with the measured energies in kcal/mol.

5. Results

Two broad classes of compounds emerged from the *in silico* screening campaign: zwitterionic (*e.g.* taurine, homotaurine, phosphoserine) and aromatic-cationic (*e.g.* 5-hydroxytryptamine). For zwitterionic compounds, the cationic moiety electrostatically interacted with the anionic glutamic acid (E) residue and the anionic moiety bound to the cationic HH region; analogously, for the cationic-aromatic compounds, the cationic segment bound to glutamate and the aromatic segment interacted with the HH region.

More specifically, optimization of taurine with Aβ in the EVHHQKLV region resulted in mostly binding interactions between the sulfonate group and K16, followed by H13. Measurable interactions at two sites within the region favoured H13-K16, with H13-H14, and E11-H14 as well. The only binding interaction with the amino group occurred with the E11-H14 orientation. The binding energies observed for taurine ranges from -616-721 kcal/mol. Homotaurine demonstrated similar results to taurine, with interactions occurring principally between the sulfonate and K16, followed by H13. Binding interactions could be measured at two sites for E11-H14 and H13-K16. As observed for taurine, the E11-H14 orientation was the only site where the amino group was bound within the EVHHQKLV region. Binding energies ranged from -620-749 kcal/mol. Analogous results were obtained in docking phosphoserine and 5-hydroxytryptamine to EVHHQKLV within Aβ (See Figure 1).

To verify these *in silico* results, *in vitro* experiments were performed confirming the capacity of homotaurine, phosphoserine and 5-hydroxytryptamine to inhibit Aβ misfolding and aggregation at low micromolar concentrations. Taurine was inactive. More importantly, the three active agents were also equally efficacious in inhibiting the misfolding and aggregation of tau and amylin peptides indicating the capacity of these endogenous agents to be evaluated against other disorders and suggesting the generalizability of this approach to a class of diseases (proteopathies) rather than single diseases or disorders.

6. Conclusions

Although administration of replacements (*e.g.*, l-thyroxine for hypothyroidism) or analogues of endogenous compounds (*e.g.*, hydrocortisone as a cortisol congener) are time-honoured therapeutic strategies, this direction has not been evaluated for AD. Searching for an "endogenous anti-AD compound" represents an unexplored concept. The utility of an *in silico* high throughput screen to aid in the identification of such endogenous compounds has been demonstrated.

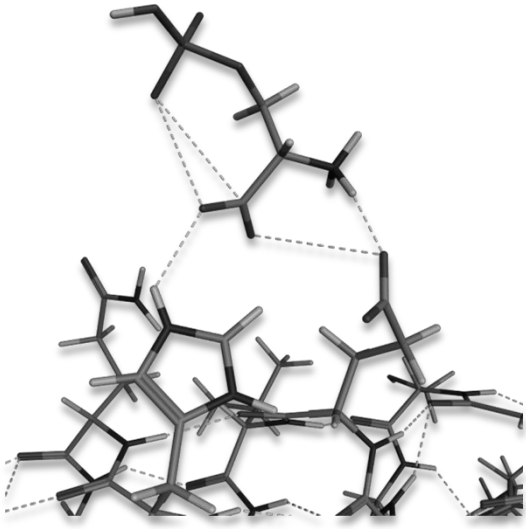

Figure 1 *The interaction between L-phosphoserine and β-amyloid at His14 and Glu11.*

References

1. Hardy J, Selkoe DJ. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* 2002; **297:** 353-356.
2. Parihar MS, Hemnani T. Alzheimer's disease pathogenesis and therapeutic interventions. *J Clin Neurosci* 2004; **11:** 456-467.
3. Giulian D, Haverkamp LJ, Yu J, *et al*. The HHQK domain of beta-amyloid provides a structural basis for the immunopathology of Alzheimer's disease. *J Biol Chem* 1998; **273:** 29719-26.
4. Stephenson VC, Heyding RA, Weaver DF. The "promiscuous drug concept" with applications to Alzheimer's disease. *FEBS Lett* 2005; **579:** 1338-42.
5. *Molecular Operating Environment (MOE)*, 2016.08; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal QC, Canada, H3A 2R7, **2017**