Principled Mechanistic Explanations in Biology:

A Case Study of Alzheimer's Disease

Sepehr Ehsani

[1] Theoretical and Philosophical Biology, Department of Philosophy, University College London, Bloomsbury, London, WC1E 6BT, United Kingdom

[2] Ronin Institute for Independent Scholarship, Montclair, NJ, 07043, United States

ehsani@uclmail.net

First version: 28 January 2021 (*Date of last revision:* 29 May 2021)

Abstract

Following an analysis of the state of investigations and clinical outcomes in the Alzheimer's research field, I argue that the widely-accepted 'amyloid cascade' mechanistic explanation of Alzheimer's disease appears to be fundamentally incomplete. In this context, I propose that a framework termed 'principled mechanism' (PM) can help with remedying this problem. First, using a series of five 'tests', PM systematically compares different components of a given mechanistic explanation against a paradigmatic set of criteria, and hints at various ways of making the mechanistic explanation more 'complete'. These steps will be demonstrated using the amyloid explanation, and its missing or problematic mechanistic elements will be highlighted. Second, PM makes an appeal for the discovery and application of 'biological principles', which approximate *ceteris paribus* laws and are operative at the level of a biological cell. As such, although thermodynamic, evolutionary, ecological and other laws or principles from chemistry and the broader life sciences could inform them, biological principles should be considered ontologically unique. These principles could augment different facets of the mechanistic explanation but also allow further *independent* nomological explanation of the phenomenon. Whilst this overall strategy can be complementary to certain 'New Mechanist' approaches, an important distinction of the PM framework is its equal attention to the explanatory utility of biological principles. Lastly, I detail two hypothetical biological principles, and show how they could each inform and improve the potentially incomplete mechanistic aspects of the amyloid explanation and also how they could provide independent explanations of the cellular features associated with Alzheimer's disease.

ACRONYMS:

Aβ amyloid beta peptide
AD Alzheimer's disease
APP amyloid precursor protein
NFT neurofibrillary tangle
NM new mechanical philosophy ('new mechanism')
PM principled mechanistic explanatory framework

Introduction

This paper aims to show the practical utility of the 'principled mechanism' (PM) account on a current case in biomedicine, namely, on the field of Alzheimer's disease (AD) research. As will be detailed throughout the paper, PM is a model of biological explanation that supplements mechanistic elements with 'principles', where the latter are thought of as non-accidental generalisations that may nevertheless fall short of full-blown lawhood. As for the subject matter, AD studies represent an active area of investigation with clear clinical and biological implications. In the philosophy of biology literature, there are (to the extent of what the research for this paper has revealed)[1] no examples of a thorough analysis of AD research; and the few AD-focused accounts that do exist concern, for instance, the idea of a "genetic cause" using the example of AD (Dekkers & Rikkert, 2006; Nordenfelt, 2006) or the subject of selfhood in AD patients (Kontos, 2004).

The paper is structured to provide a philosophical case study of the disease, centred on three questions: (i) what is the current problem with AD research and why do we face this problem? (**Section 2**); (ii) why might contemporary philosophical accounts of mechanism in biology not offer tractable solutions to adequately confront this challenge? (**Section 3**); and, (iii) what can the PM framework offer instead to move in a more productive direction (**Sections 4–6**)?

The AD field is broad and varied, and I have tried, to the extent possible, to sufficiently represent the current breadth of research whilst also staying on point. I aim to explicate the following answers to the above questions: First, the AD field is, for the most part, of one voice when it comes to the central elements of a mechanistic explanation that has motivated research in the field for the past three decades. The AD mechanistic model and resultant explanation are greatly detailed (and increasingly so by the day). However, the explanation — despite immense effort — has not translated into the clinic for patients, and many clinical trials have been unsuccessful. I argue that this is not due to challenges in clinical 'translations' of the explanation's predictions, but rather to a fundamental problem with the way in which mechanistic explanations are approached in the field.

Second, I argue that the solution to this problem is not to simply do away with mechanistic explanations, for there is no serious replacement for this explanatory type in biomedicine. The solution should build on the existing mechanistic explanations. Moreover, whilst sympathetic to the "New Mechanist" (NM) project in the philosophy of scientific explanation, I provide several strands of argument to the effect that existing NM accounts, to the extent that they go beyond reflecting current biological practice, may not be sufficiently critical of that practice to be used to help resolve the problem set out in this paper. Third, I will propose that a series of 'paradigmatic' tests can, as an initial strategy, point to elements within a given mechanistic explanation that could make it more comprehensive and generate new empirical questions. As a next step, I will argue that the addition of biological principles that could govern or apply to the elements highlighted in the tests can make the overall explanation richer, and the potential understanding of AD pathobiology more fundamental.

[1] Of note, a relevant recent philosophical work presents a framework called 'MecCog' that builds mechanism schemas for diseases including AD (Kundu, Darden, & Moult, 2020).

Before proceeding further, I will address two important concerns in **Section 1**, namely, what range of mechanistic concepts the paper makes use of, and why AD is well-suited for the purpose of this case study.

1. Usage of mechanistic concepts, and context of AD research in biology

Perusing current biomedical journals would leave little doubt that the search for 'mechanisms' forms the basis of a great proportion of the research effort. However, there is much ambiguity in the intended meaning of mechanistic concepts and terms in the biological literature (Marder, 2020). This might partly be attributable to the impression that 'mechanism' could be said to be taken as a *primitive* concept in biology (much like 'point' or 'line' in geometry (Pearl & Mackenzie, 2018, p. 373)) — a notion on which much else relies, but one which is refractory to a simple and universal definition. It goes without saying that philosophers of mechanism have studied these topics extensively, but this literature has not yet found its way into mainstream biomedical research. For our purposes here, however, the intended terminological meanings should be crystallised.

1.1. Actual mechanisms, mechanistic models and mechanistic explanations

Consider the statement "the bacterium *Helicobacter pylori* can lead to (or cause) peptic ulcers". Here, 'peptic ulcer' is the *phenomenon to be explained* (henceforth the 'phenomenon'). Moreover, *H. pylori* is a putative cause, or a causative agent. One could then ask why *H. pylori* causes peptic ulcers. Perhaps because this is its function in the gastrointestinal environment. The question could be rephrased as: how or by what means does *H. pylori* cause peptic ulcers? The 'how' or the 'means' with which a putative causative agent leads to a phenomenon can be called the *actual mechanism* by which the phenomenon arises (henceforth the 'mechanism').[2] To avoid confusion, I take the word 'mechanism' to indicate our assumption about the work that causes or produces a phenomenon.[3]

Now, we can model a mechanism. One can take a model to be some representation of reality (or a phenomenon), and it can be used as a heuristic or thinking tool, or as part of a broader explanation, or perhaps as a way to simplify the complexity of the phenomenon. There are different accounts of how a model does the 'representing', but I will not expand on them here.[4] In molecular biology, cellular processes are often represented using a network of parts (e.g. proteins, genes, RNAs, etc.), indications of change and movement (e.g. of a protein from one location to another), interactions (e.g. protein–protein, protein–DNA, protein–lipid), and so on. Therefore, we can define a *mechanistic model* as a representation of a mechanism. It follows that a given mechanism could have multiple different representations.

[2] There are distinctions to be made between mechanistic and 'difference-making' accounts of causation, but I will not delve into the topic here.
[3] The reason why this assumption may not apply across the board is that in some cases using 'mechanism' as a stand-in for 'means' may strike as odd: e.g. when referring to the means with which a magnetic pole attracts a magnet, talking about 'mechanism' is unusual; rather, we could talk of magnetic 'fields' and 'forces'.
[4] As an example, a candidate for how models represent has been called the 'DEKI' account, entailing four elements: denotation, exemplification, keying up of properties, and imputation (Frigg & Nguyen, 2016).

Mechanistic models can subsequently be used as part of a statement or description to detail and give reasons for a phenomenon (see e.g. (Brini, Simmerling, & Dill, 2020)). Let us call this the *mechanistic account* or *mechanistic explanation*. A mechanistic model could give rise to several explanations, which can be used by investigators for predictions, molecular interventions, etc. (Lombrozo, 2011).

Three caveats are necessary: First, some studies in cellular biology may stop at the stage of detailing certain elements of a model, and not progress to a full-fledged mechanistic explanation. For our purposes here, however, the explanatory stage is key. Second, biologists investigating a cellular phenomenon may at times suffice to provide an analysis of the putative function of a 'part', such as a protein, and only provide a version of a 'functional' explanation. As will also be seen later in the context of AD, this should indeed be viewed as complementary to mechanistic explanations (see also (Theurer, 2018)). And third, I am not claiming that all explanations require a model, but that mechanistic explanations typically use at least some elements of mechanistic models.

1.2. AD research as a quintessential mechanistic research programme

AD poses a huge burden on patients and their families. Worldwide, the projected number of dementia cases (a majority of which is thought to be AD) is estimated to reach over 131.5 million affected individuals by 2050 (Editors, 2016). Why is AD a suitable research field to study the theoretical basis behind mechanistic explanations? First, there has now been more than 50 years of systematic research on AD, and a vast portion of the published works refer to "disease mechanism", "mechanistic understanding", "mechanistic pathways" and many other related concepts. The current cellular understanding of AD is complex to the point that a lot can be said about its various facets, its strengths and its shortcomings. Furthermore, a well-accepted and overarching mechanistic explanation has been the mainstay of the field for several decades, and, importantly, it has been tested in different lights, i.e. in countless laboratory-based assays all the way to many clinical trials. So, what exactly is the problem?

2. Problems for the mainstream biomedical research approach to AD

In 1907, Alois Alzheimer (1864–1915) provided two pathological hallmarks for the disease in the brain, which are now widely known as *amyloid plaques* and *neurofibrillary tangles* (Alzheimer, Stelzmann, Schnitzlein, & Murtagh, 1995) (these two terms will be referred to repeatedly). The hallmarks are now known to be due to the aggregation of two main types of 'sticky' proteins: the plaques are formed of *amyloid beta (Aβ) peptide* (mostly outside affected neurons), and the neurofibrillary tangles (NFTs) of *tau protein* (inside affected neurons). Focused investigations of Aβ three decades ago are what started the extensive and what one could call the 'mainstream' molecular research effort on AD. What has resulted is a working explanation called the *amyloid cascade hypothesis/mechanistic explanation* (Beyreuther & Masters, 1991; Hardy & Allsop, 1991; Selkoe, 1991), henceforth the 'amyloid explanation'. Although many additional genes and cellular pathways have since been associated to varying degrees with some manifestations of the disease (see (Liu, Xie, Meng, & Kang, 2019)), almost all findings in AD are interpreted via and/or placed within the amyloid explanation.

2.1. The amyloid cascade mechanistic explanation

To analyse the amyloid explanation, it would help to choose an illustrative schematic that depicts both processes of amyloid plaque formation by Aβ and NFT formation by tau. However, note that, technically, the amyloid cascade hypothesis usually only refers to the amyloid plaque formation process in AD and does not include the NFT formation arm. Be that as it may, to simplify subsequent references to these two defining lesions of AD, I will use the umbrella term of 'amyloid explanation' to refer to both processes.

Of the available choices, a summarising figure from a review article by Panza and colleagues (Panza, Lozupone, Logroscino, & Imbimbo, 2019) is shown in **Figure 1** on the next page and will be used as a reference visual for the amyloid mechanistic model. Also, to simplify the depicted illustration using text, three salient processes in the figure are summarised in **Table 1**. These are:

(i) The *amyloidogenic* amyloid precursor protein (APP) processing pathway; this process *increases* the production of Aβ and the formation of amyloid plaque.

(ii) The *non-amyloidogenic* APP processing pathway; this process *decreases* the production of Aβ and the formation of amyloid plaque.

(iii) The NFT formation pathway; this process leads to the aggregation of tau.

In the figure, the lipid membrane of a neuronal cell is shown as a bilayer (two lines of circles with lines in between, like a horizontal ladder). The cytoplasm or intracellular space is the area below the membrane, shown in beige. All the area above the membrane is the extracellular space, and one more neuron and two other cell types in the brain (astrocytes and microglial immune cells, which act to support neurons) are illustrated at a much smaller size. Pathways (i) and (ii), which have to do with amyloid plaque formation, both start with the APP protein, which in the figure is shown as a twisting 'tube' that has been cut by two pairs of scissors. The reason for the three colours of the tube (red, orange and blue) is to show the three fragments (or 'peptides') of APP that form when it is cleaved by a number of other proteins (enzymes).

In (i) the *amyloidogenic pathway*, APP is cleaved by an enzyme ('β-secretase' or BACE, depicted as a pair of purple scissors above the membrane) to produce two fragments, namely 'sAPPβ' (in blue) and 'C99' (the part of APP that remains). The C99 fragment is then cleaved by another enzyme, 'γ-secretase' (depicted as a pair of yellow scissors within the membrane), yielding the Aβ peptide (orange fragment). Aβ 'monomers' start to aggregate to form oligomers, fibrils and plaques (shown sequentially in the extracellular space). Over the past decade, attention has shifted from the insoluble fibrils and plaques to the soluble oligomers as being the toxic Aβ form in neurons (Dear et al., 2020). These steps are summarised in the first column of **Table 1**.

Concomitant with (i), the (ii) *non-amyloidogenic pathway* is also proposed to be taking place on the cell membrane with respect to APP. Specifically, in this pathway, rather than initially being cleaved by β-secretase, APP is cleaved by 'α-secretase' at a different site, leading to the production of a soluble fragment called 'sAPPα' and *preventing* Aβ peptide production. Because this pathway is supposed to be a 'physiologic' process that happens normally in the brain, it is not shown in the figure. However, the steps are summarised

in the second column of **Table 1**. In AD, the *balance* between these two pathways is thought to be skewed toward the amyloidogenic route. In other words, more of process (i) and less of process (ii) take place.

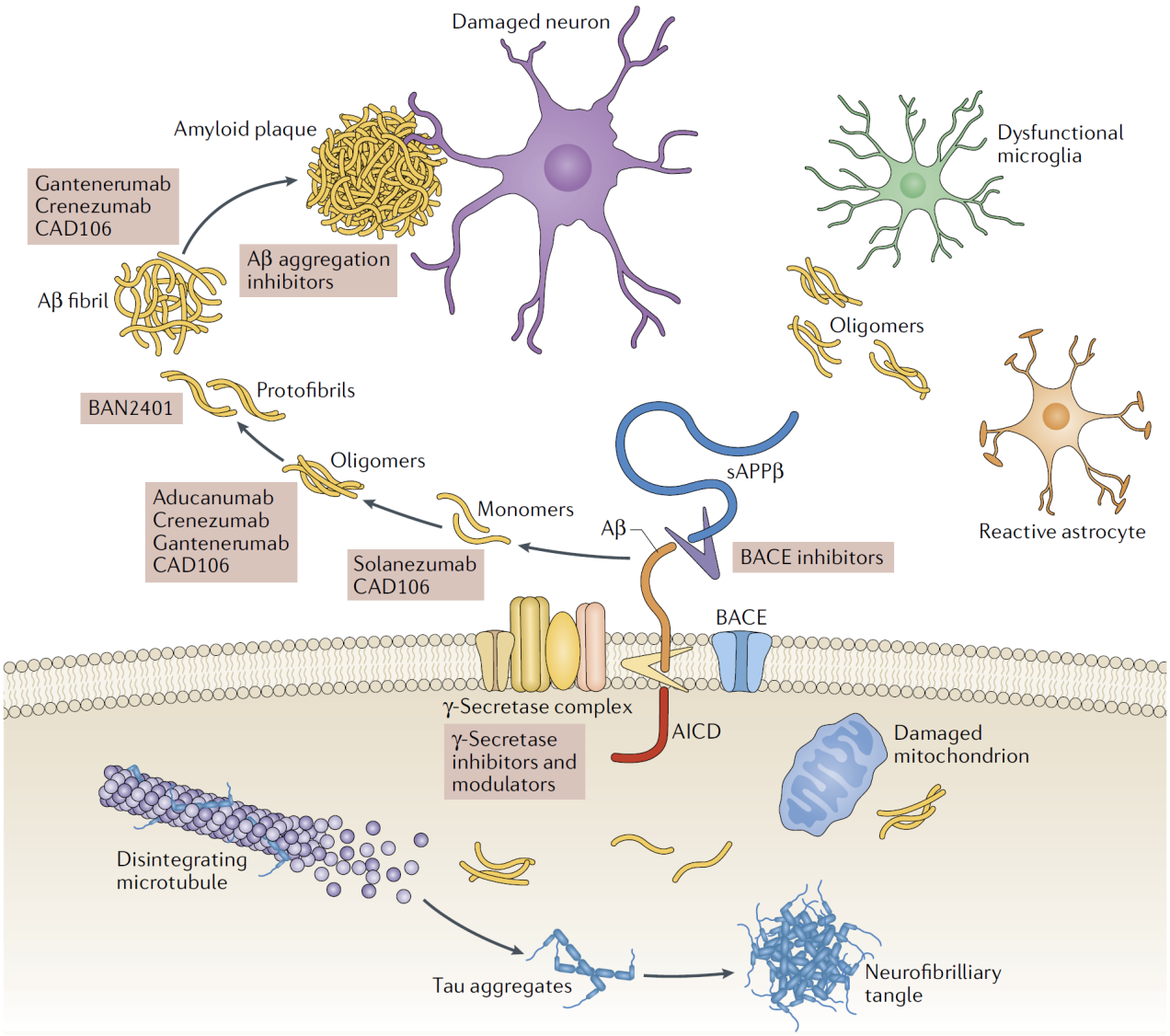

Figure 1. Putative pathobiological mechanistic model of Alzheimer's disease (Panza et al., 2019). Here the authors show the process of amyloid plaque formation outside the neuronal cell, and the generation of neurofibrillary tangles inside the cell, along with "mechanisms of action of the main anti-amyloid-β (Aβ) drugs that are currently in phase III clinical development for the treatment of Alzheimer disease". The intracellular endpoint depicted is mitochondrial damage, whereas the extracellular endpoints shown are neuronal damage, microglial dysfunction and reactive astrocytes. Abbreviations, as per the original caption, are: AICD, amyloid precursor protein intracellular domain; BACE, β-secretase; sAPPβ, soluble amyloid precursor protein-β. Reproduced with permission from Springer Nature.

Pathway (iii) concerns the intracellular formation of NFT, where the central player is the tau protein. Tau is mainly found in the axons (long projections) of neurons and is physiologically associated with promoting the assembly of microtubules. Microtubules are polymers of 'tubulin' proteins and form part of the cytoskeleton: they provide structure to the cell and aid in other functions. In **Figure 1**, a microtubule is depicted as a purple cylinder (formed of many tubulin 'spheres') below the membrane, and tau proteins can be seen as blue objects

wrapping around the microtubule. In the AD amyloid explanation (and indeed other neurodegenerative conditions involving tau), the tau protein, which is modified in the cell by the attachment of a 'phosphoryl' chemical group (i.e. is said to be 'phosphorylated') on a number of its amino acids, becomes *hyper*phosphorylated, indicating that multiple sites on the protein become saturated with phosphorylation. It is thought that tau hyperphosphorylation leads to its dissociation from microtubules, in turn leading to microtubule disintegration and the formation of tau aggregates (both depicted in the figure). These aggregates can turn into NFTs and may even spread to other neurons (not shown in the figure), causing further NFTs to form in a given brain region. The steps in this pathway are summarised in the last column of the table.

Table 1. Three pathways in the AD pathogenesis mechanism (*details adapted from* (Panza et al., 2019))[5]

Amyloidogenic APP Processing	**Non-Amyloidogenic APP Processing**	**Neurofibrillary Tangle Formation**
Increased in AD	*Reduced* in AD	*Increased* in AD
START ○ Amyloid Precursor Protein (APP) **PROCESS** 1. Cleavage by protease β-secretase 1 (BACE1) **PRODUCTS** ○ Soluble extracellular fragment, sAPPβ ○ Cell-membrane-bound fragment, C99 **PROCESS** 2. Cleavage of C99 by protease γ-secretase **PRODUCTS** ○ Release of Aβ peptide (40- or 42-amino-acid length) ○ Aggregation of Aβ to form oligomers, protofibrils, fibrils and plaques	**START** ○ Amyloid Precursor Protein (APP) **PROCESS** 1. Cleavage by extracellular protease α-secretase **PRODUCT** ○ Soluble extracellular fragment, sAPPα	**START** ○ Microtubule-associated tau protein **PROCESS** 1. Hyperphosphorylation of tau **PRODUCTS** ○ Aggregation of hyperphosphorylated tau ○ NFT formation
ENDPOINT ○ Synapse loss and neuronal death mediated by Aβ oligomers		**ENDPOINTS** ○ Cytoskeletal changes and disruption of axonal transport ○ Spread of misfolded/aggregated tau to other cells

[5] The steps in **Table 1** are extensively simplified, given that for each step multiple details and subtleties have been reported. One such important detail, for example, is the varying clearance of Aβ by 'apolipoprotein E' (APOE) proteins (Yamazaki, Zhao, Caulfield, Liu, & Bu, 2019). In fact, variants of APOE are genetic risk factors for late-onset sporadic AD (Ishii & Iadecola, 2020).

The amyloid explanation revolves around these protein aggregation events, with the end result being neuronal loss and cognitive impairment. What happens in the brain at a cortical level (i.e. populations of neurons and brain regions) is beyond the current discussion, as we are concerned with the more immediate task of accounting for processes at the cellular level. Nonetheless, attention in the field has been directed for more than a decade now toward a region of the brain called the entorhinal cortex as perhaps the initial part of the brain suffering from the loss of its neurons (NTNU, 2020; Shugart, 2020).

The elucidation of each of the mentioned three pathways has been a major advance in the field, and for instance in the case of tau aggregation, research on its mechanistic details has been featured on the cover of *Nature* as recently as November 2019. This shows the importance with which deciphering AD's pathobiology has been judged not just by the AD research field but also by the broader scientific community.

What might now be apparent is that a conceptual gap that the amyloid explanation needs to address is, for example, how pathways i and ii interact and influence pathway iii. The nonamyloidogenic pathway is supposedly a normal physiological process, whereas the amyloidogenic pathway is pronounced in disease. Yet, to a certain extent, both are taking place in health *and* disease. Moreover, pathway iii (involving tau aggregates) has been identified to have non-neurodegenerative manifestations as well (Park et al., 2020), and to also participate in normal physiologic processes in, for example, hibernating animals (Arendt et al., 2003). The picture is therefore complicated in various ways. Having said that, this is not a complication that would, at least conceptually, pose a problem for current frameworks within the philosophy of biology to address. For us, the nub of the problem is more fundamental than how the various pathways may interact.

2.2. Capturing the full complexity of the disease

Up to this day, the amyloid explanation has actually been quite *successful* in a certain sense: details of various aspects of the underlying working model continue to be worked out by different groups, clinical trials are designed based on the explanation, and the 'tenet' of the explanation (i.e. the theme of protein aggregation inside and outside neurons) remains quite intact. As mentioned earlier, even when alternative explanations are proposed for AD, they are often pinned to the amyloid explanation. As a case in point, there has been an ongoing interest in the possibility of a microbial link to AD (Cairns et al., 2020), and whether Aβ has antimicrobial properties (Pastore, Raimondi, Rajendran, & Temussi, 2020). But when an "antimicrobial protection hypothesis" was proposed to account for these possibilities, the authors commented that "the new model extends but remains broadly consistent with the Amyloid Cascade Hypothesis and overwhelming data showing the primacy of Aβ in AD pathology" (Moir, Lathe, & Tanzi, 2018, p. 1602). It might be safe to say that the amyloid explanation has, in a way, stood the test of time. Furthermore, it is hard to imagine a *radically different* mechanistic explanation that could circumvent the current one, and the AD hallmarks, altogether.

Nonetheless, there is in fact a fundamental and consequential problem with the amyloid explanation, and it is that the explanation has not led to genuine clinical improvement in patients. This reality is highlighted when we consider the many clinical trials that have tested what were meant to be 'disease-modifying' therapeutics validated in the laboratory based on the amyloid account. In an opinion piece titled "we need to

radically rethink our approach to Alzheimer's research", Mark Hammond and Tim Newton write that "over the past decade we've seen failure after failure in clinical trials for neurodegenerative disease [and] despite over 200 clinical trials, we still don't have any meaningful therapeutics for Alzheimer's" (Hammond & Newton, 2020). This does not stem from the way in which the mechanistic explanation is being *translated* from 'bench to bedside'. Practicable strategies already exist in biology to deal with the problem of translation (see e.g. (Henderson, Rieder, & Wynn, 2020)).

In fact, many of the AD therapies tested have been produced to *match* the understanding and prediction afforded by the amyloid explanation in remarkably precise ways. In other words, if the explanation suggests that protein 'X' should be lowered in patients (and this has been borne out in laboratory and animal experiments), the therapeutics being tested *are* indeed lowering that protein in patients being tested. Hence, the explanation's predictions/demands are being *translated* all the way to patients in a strictly biological sense but are still not successful in halting AD. Something might therefore be amiss with the explanation itself, in that it is not capturing the full complexity of AD and is somehow seriously incomplete.

It is important to note here that such problems could in fact generalise to — and are typical of — any other mechanistic explanation in biomedicine. In cases such as certain rare monogenic diseases, the explanation might have fewer identified components relative to the amyloid explanation of AD and a more delineated causal chain, whereas in other complex cases such as various types of cancer the mechanistic explanation might be much more multifaceted than that of AD. Furthermore, there might also not be a general consensus in the respective research field on a unified mechanistic explanation. That being said, this paper's analysis could just as easily be applied to the mechanistic explanation of any pathobiological condition.

Now, how should the challenge with the amyloid explanation be approached? I will outline two possibilities, calling them the 'incremental' and the 'non-mechanistic' options, and then point out that my thesis advocates for a middle ground. To begin with, solutions proposed by some AD investigators appear to argue for staying within the confines of the current mechanistic framework, whilst attempting to increase our understanding of certain details and gaps about the workings of the model, or the timeline of clinical trials and interventions (Aisen, 2019; Petsko, 2018). I call this the incremental approach. An important detail requiring resolution, for instance, concerns Aβ, such as clarifying how its levels in patients are correlated with the stage of the disease (Masters, 2019), or whether it is causative of AD or a byproduct (Panza et al., 2019).

The focus on details might also lead to the addition of new molecules to the explanation, which may eventually become therapeutic 'targets'. However, given the long history of AD research, it is quite unlikely that one or more hitherto undiscovered targets waiting to be added to the mechanistic model would suddenly change everything. Nevertheless, I cannot 'prove' this, for such an outcome is in theory possible, although, as I have emphasised, all new details that have thus far been discovered have revolved around the amyloid explanation.

The incremental approach might also reveal new knowledge about the existing molecular parts of the amyloid explanation. This is important because the evolutionary history behind the function of each protein

stretches back hundreds of millions of years, and any key protein identified in the AD explanation (e.g. APP) very likely also functions in various other cellular processes that might have little to do with AD per se, therefore making interventions on it risky. Additionally, when investigators focus on one mechanism, a general problem of 'masking' may be encountered, whereby "the operation of one mechanism might *mask* or hide the operation" of other mechanisms (Illari, 2011, p. 146). For example, the secretase enzymes might function in as-yet-unidentified processes in neurons that are critical to normal cognition. If that is so, the amyloid explanation must be able to capture these important aspects of its key protein players as well. However, there is no obvious path as to how the amyloid explanation can connect with other cellular mechanisms. The incremental approach can in principle reveal insights toward this issue, but it is not clear how exactly this would be achieved.

An opposing option relative to the one above is to look for completely non-mechanistic explanations for AD, but such a suggestion would be received as outlandish in cellular biology. It is inconceivable for anyone in biomedicine to advocate giving up the entire 'mechanistic enterprise' (i.e. the creation of mechanistic models and provision of mechanistic explanations), for there is no viable alternative to such an approach. If someone proposed a completely non-mechanistic approach to studying AD's cellular phenomena — it is hard to imagine what this would look like, for even a purely mathematical approach would have to rely upon some 'platform' in the form of a mechanistic model — it just cannot realistically inform current research, as it would simply not 'connect with' anything that investigators are pursuing. Consider also that even network-based or topological explanations in molecular biology are often used to search for mechanistic elements within them (see (Yadav, Vidal, & Luck, 2020)). What is more, the mechanistic enterprise appears to connect with at least some of our intuitive insights about how nature works (Spelke & Kinzler, 2007).

My thesis argues for a middle way, such that the meticulous protein-aggregation-centred mechanistic understanding is preserved but also augmented with a 'novel' non-mechanistic approach so that new experiments would not merely bring about 'more of the same', but rather genuine steps forward in our understanding of the complexity of the disease. This is the ultimate aim of the PM framework, to help show how exactly the amyloid explanation might be 'incomplete', and how one might go about improving it. One can then reasonably hope that the explanation would have a better chance of actually halting the disease.

But first, before outlining the PM approach, I will look for insights toward improving the amyloid explanation in work by contemporary philosophers concerned with mechanistic theory.

3. 'New Mechanism' and clues for moving past the explanatory problem

Beginning in the 1990s, a renewed interest in mechanisms began to take shape primarily within the philosophy of biology. This strand of investigation is termed "New Mechanism" (NM) to distinguish it from research on the history of the Mechanical Philosophy. NM was a welcome development, for it aimed to systematise — and provide the underlying theory for — mechanistic concepts and usages that were (and for the most part still are) treated superficially in biomedical practice.

A reminder of the context within which NM arose is important here. Many philosophers of science in the mid-twentieth century were pursuing a nomological (i.e. natural-law-based) tradition, best exemplified in Carl Hempel's (1905–1997) deductive-nomological (DN) model of scientific explanation (Hempel, 1965). The zeitgeist of this approach can perhaps be said to have been "experiment as the source of knowledge, mathematical formulation as the descriptive medium [and] mathematical deduction as the guiding principle in the search for new phenomena to be verified by experimentation" (Dijksterhuis, 1961, p. 3). This also instilled a sense of valuing "rigor, precision, and generality" in scientific theorising (Bogen, 2020).

But the DN approach was problematic, even in the seemingly compatible subfields of physics. Already in the same period as the peak of logical empiricism, Thomas Kuhn (1922–1996) was questioning the highly formal and axiomatic approaches to scientific theories (Kuhn, 1962), and, in the words of Simon Blackburn, advocated a "less formal and more contextualized approach […which] stressed the open-endedness of scientific activity [and] the heuristic value of analogies and models" (Blackburn, 2016, p. 475). Philosophers of biology also took note of such problems. William Wimsatt, who together with a group of students and collaborators set the stage for the NM approach, writes that in 1974, he argued that "discovering a mechanism as a relatively stable and manipulable articulation of causal factors better fit the activity of biologists than a search for laws" (Wimsatt, 2017, p. xv).

Over the past 30 years, NM philosophers have provided influential accounts with which to systematically conceive of mechanisms and mechanistic explanations. For example, in 1993, William Bechtel and Robert Richardson wrote that mechanistic research results in "a detailed account of the parts and operations of a mechanism and how they are organized and orchestrated in a specific model system" (Bechtel & Richardson, 2010, p. xli)[6]. Peter Machamer, Lindley Darden and Carl Craver defined mechanisms as "entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions" (Machamer, Darden, & Craver, 2000, p. 3). Along the same lines, Bechtel and Adele Abrahamsen defined a mechanism as "a structure performing a function in virtue of its component parts, component operations, and their organization [where] The orchestrated functioning of the mechanism is responsible for one or more phenomena" (Bechtel & Abrahamsen, 2005, p. 423). Finally, Stuart Glennan and Phyllis Illari have written about the concept of a "minimal mechanism", whereby "a mechanism for a phenomenon consists of entities (or parts) whose activities and interactions are organized so as to be responsible for the phenomenon" (Glennan, 2017, p. 17; Glennan & Illari, 2017, p. 92).

Without going into the specifics, the way in which biologists think about the AD mechanism arguably matches the accounts above. Moreover, the amyloid explanation itself is attuned to the way that, for example, Bechtel and Richardson defined mechanistic accounts. However, it is not immediately clear based on these accounts how the amyloid explanation might be incomplete and how it could be improved. Depending on the context, the NM accounts require much more expansion to function as a critiquing framework. One could, for example, 'read into' the accounts or find elements that could be useful for our problem, but they do not provide explicit details for us to utilise. Part of the reason for this might be that many NM philosophers and philosophers of biology in general might not have concentrated on the *failures* present in modern cell biology. Laura Franklin-

[6] This is a reprint of the 1993 edition. The quoted text is from the introduction to the original edition.

Hall similarly suggests that “a too-successful enculturation of philosophers into the scientific mindset [makes] it difficult to achieve the critical distance needed to philosophize *about* science” (Franklin-Hall, 2016, p. 71).

Notwithstanding these concerns, what *hints* could one distil from the NM approach as to what the amyloid explanation might be lacking? I believe such hints might become apparent when one considers how the NM and related approaches have dealt with ideas concerning the introduction of laws into mechanistic explanations. Craver and James Tabery maintain that mechanisms “seem to play the role of laws in the biological sciences: we seek mechanisms to explain, predict, and control phenomena in nature even if mechanisms lack many of the characteristics definitive of laws in the logical empiricist framework (such as universality, inviolable necessity, or unrestricted scope)” (Craver & Tabery, 2019). In another take, Craver and Marie Kaiser write that “mechanists *decenter* laws in their thinking about science because the old paradigm, centering laws, has become mired in debates that are inconsequential and, as a result, have *stopped generating new questions* and producing new results” [*emphasis added*] (Craver & Kaiser, 2013, p. 144). Whilst a goal of my thesis is to put laws back in the centre of discussion,[7] I share the view of Craver and Kaiser in emphasising the importance of generating new empirical questions as a key criterion of a framework’s success. Let us call this Hint #1.

Now, more compatible with the view on laws argued in this paper is that of Bechtel and Richardson, who also considered the place of laws or “general and abstract explanatory principles” (Bechtel & Richardson, 2010, p. 232) in mechanistic accounts. Their main motivation here was to “question the hegemony of laws in explanation [but] not their existence” (Bechtel & Richardson, 2010, p. 256). In more recent work, Bechtel comments that “laws may be invoked to characterize the overall functioning of the mechanism or some of its operations, but it is the discovery that particular operations are being performed that is required to specify the mechanism” (Bechtel, 2011, p. 537). Congruent to this description, Nancy Cartwright, John Pemberton and Sarah Wieten (whose work is not usually considered part of the NM tradition) have proposed that “when a mechanism M gives rise to a regular behavior RB that is described in a cp [*ceteris paribus*] law, RB is what it takes for some set of principles that govern the features of M’s parts in their arrangement in M all to be instanced together” (Cartwright, Pemberton, & Wieten, 2020, p. 18). I will say more on *ceteris paribus* laws in **Section 6**.

In recent work analysing examples from chronobiology (the biology of time in terms of an organism’s 24-hour circadian rhythm), Bechtel mentions “principles of organization (design principles)” which “assert that any system implementing the organization will exhibit the specified behavior” (Bechtel, 2017, p. 19). Relatedly, Sara Green has investigated the idea of design principles in the context of ‘systems biology’ and introduced the idea of “constraint-based generality” (Green, 2015). Whilst my thesis is sympathetic to these accounts, again they do not explicitly show how the amyloid explanation should be overhauled, and the role and discovery of principles come across as peripheral to the specification and operations of a mechanism. I state this because, for instance, an AD investigator might just assume that the current amyloid explanation is *already* serving the purpose of ‘organising’ the components of the underlying model by accounting for the temporal

[7] Others, e.g. Bert Leuridan in (Leuridan, 2010), have also argued that laws cannot be supplanted by mechanisms in scientific explanations.

cascade of events both inside and outside the cell. Thus, the investigator needs a convincing reason to invoke an extra concept, such as a 'principle'. Nonetheless, the NM works cited in this and the preceding paragraph hint at an explanatory role that principles of organisation could play, at least in a qualitative sense. Let us consider this Hint #2.

As might be expected from the often-cyclical nature of the history of ideas, as the NM approach began to become dominant in the philosophy of biology, there was a small but steady wave of proposals that reintroduced or amalgamated certain elements of the pre-NM nomological approaches. For example, José Díez has proposed a 'neo-Hempelian' account of scientific explanation, whereby "to explain a phenomenon is to make it expectable by introducing new conceptual/ontological machinery and using special, and non-ad hoc, non-accidental regularities" (Díez, 2014, p. 1413). A key notion in Díez's account is that of 'expectability'. Roger Deulofeu and Javier Suárez pick up on this notion in their paper on "when mechanisms are not enough", and write that in their analysis, "the use of scientific laws is supposed to be a minimal *requirement* of all scientific explanations, since the purpose of a scientific explanation is to make phenomena expectable" (Deulofeu & Suárez, 2018, p. 95) [*emphasis added*].

In this body of literature, one of the closest approaches to the one in this paper is an account by Karina Alleva, Díez and Lucia Federico on a problem in biochemistry (Alleva, Diez, & Federico, 2017). The authors are concerned with the mechanistic description of the conformational changes in a protein upon the binding of another molecule ('allosteric' regulation). What they propose "essentially contains nonaccidental, nomological regularities that can properly be considered as laws in a relevant, though minimal, sense of lawhood" (Alleva et al., 2017, p. 12). Importantly, they say that "we do not believe that the mechanistic and our model-theoretic accounts are in opposition" (Alleva et al., 2017, p. 12), but rather "we advocate a plural, syncretic perspective in which every relevant aspect is explicated according to its specific nature" (Alleva et al., 2017, p. 13). This "plural, syncretic perspective" and the notion of 'expectability', even though originated in works which one might say are 'reacting' to the NM tradition, could still be our Hint #3.

The confluence of these three hints, i.e. a framework's ability to generate new empirical questions, an explanatory role for principles, and an explanation's drive toward expectability, is a niche where the 'principled mechanistic' (PM) explanation of this paper can fit. As I shall argue, PM's strength lies first and foremost in generating questions that current mechanistic explanations are not poised to produce. Furthermore, I will argue for an independent and critical explanatory role for 'biological principles'. And, lastly, PM pushes for *quantitative* expectability of the phenomenon being studied. More on these in the next sections.

Beyond the above hints, there are also other intersections between PM and NM. For instance, part of the PM framework involves creating a series of 'tests' to detail what a 'paradigmatic' PM explanation would look like (**Section 4**). For this, concepts developed in the NM literature about mechanistic elements such as decomposable parts, organisation and levels will be of great utility. This is why I consider the PM project to be complementary to NM, yet more prescriptive of what investigators 'ought' to be doing compared to NM when it comes to existing problems in biology.

In closing this section, I should emphasise that my intention is not to imply that all open problems related to current AD research require PM, or some other augmented mechanistic framework. The purpose of PM is to improve the overall explanation of AD. Nonetheless, recalling the final paragraph of **Subsection 2.1**, I noted that there are open challenges pertaining to the amyloid explanation that are separate from what I argue to be its 'foundational' shortcomings. One such challenge had to do with how one could reconcile the involvement of pathways i, ii and iii in both physiologic and disease processes. Would, for instance, disease in such cases be a matter of imbalance? Accounts in NM, such as those explicating how operations within a mechanism might be organised or what different levels and feedback might exist in a particular mechanistic explanation of a phenomenon, would be well-placed to tackle such questions.

4. Toward an ideal 'principled mechanistic' explanation in the context of AD

The paper thus far has indicated that the PM approach aims to strike a middle way between purely mechanistic and non-mechanistic options. Taking a cue from NM philosophers, on the mechanistic side, PM aims for some systematisation of the elements of mechanistic explanations. On the non-mechanistic side, the goal is to adapt hypothesised 'biological principles' to mechanistic explanations, where the resultant PM explanation would be an augmented but still a coherent whole.

To start to demonstrate that satisfactory explanations in biology are liable to require appeal to principles as well as mechanisms, it might be useful to mention an analogy with 'reaction mechanisms' and thermodynamic laws in chemistry. When a chemist sets out to explain the reaction between some molecules, they detail the reaction mechanism, i.e. the sequence of molecular/atomic steps leading to an overall chemical change. But they do not stop there, for they also appeal to thermodynamic laws that 'govern' the steps.

In fact, thermodynamics (e.g. the second law) has been highly formalised in chemistry in the form of the notion of change in 'Gibbs free energy' (ΔG), "whose sign predicts the direction of reaction, and whose magnitude indicates the maximum amount of work realizable from the reaction" (Feinman & Fine, 2004, p. 2). As such, thermodynamics can help (i) *predict* something about the reaction (recall the notion of 'expectability' from Hint #3 in the previous section) and (ii) provide some sort of *quantification* for the yield of the reaction. These could arguably be achieved without necessarily appealing to the mechanistic details of the reaction. However, the notion of ΔG can also (iii) be applied to the details of the reaction mechanism itself to reveal new insights. As an example, chemists have used thermodynamic notions of molecules' lowest free-energy states to study the immensely complex mechanism of hydrogen bond formation amongst clusters of water molecules (Richardson et al., 2016), allowing them to reveal how the bond formation mechanism *itself* might be operating (see also (Llored, 2011)).

In today's cell biology, we have the equivalent of intricate 'reaction mechanisms' for phenomena such as AD (although perhaps not as systematised as in chemistry), but what is largely missing is the cell biological equivalent of thermodynamic laws. This, in essence, is the framework that a PM explanation aims to achieve. The tasks now at hand are to describe a way for systematising mechanistic explanations, and to show how an

appeal to principles can augment the power of cell biological explanations. This section embarks on the first task, with **Sections 5** and **6** taking up the latter lead.

Here, based on different NM-related accounts (Bechtel & Abrahamsen, 2005; Bechtel & Richardson, 2010; Cartwright et al., 2020; Glennan & Illari, 2017), earlier approaches to naturalistic explanations, and also research on AD and non-AD mechanistic explanations in the cell biology literature, I propose to develop a concise series of criteria for a *paradigmatically* 'good' mechanistic explanation. These criteria could be thought of as candidate determinants of the 'good-making features' (borrowing from (Newton-Smith, 1981)) of a mechanistic explanation. The idea is not to say that fulfilling all the criteria would arrive at an ideal or complete explanation *in an absolute sense*, but I only claim that moving toward fulfilling these criteria, if we can, would produce better mechanistic explanations than what we already have. Furthermore, the criteria can be presented in the format of a series of 'tests', simply to help with determining whether an explanation meets the criteria.

- **TEST #1:** The phenomenon to be explained is set out as unambiguously as possible.
  This first test concerns the explanandum. It is only referring to the phenomenon and is not yet concerned with the explanation per se, and is simply appealing to a goal of clear stipulation of that which a biologist is to investigate. There is often much work to be done prior to setting out to explain a phenomenon in clarifying the biological problem at hand. Such an aim not only helps to avoid explanatory irrelevance but could also facilitate consistency amongst investigators and the ease of communicating different explanatory angles of the phenomenon. For example, setting out to explain 'cell death' (e.g. apoptosis) is clearer than setting out to explain 'cellular dysfunction', without stipulating what the 'dysfunction' implies. In a sense, an investigator wants to pre-empt the question: 'but what exactly *are* we explaining?'.

- **TEST #2:** The explanation sets out an environment to situate the mechanism leading to the phenomenon and refers to decomposable and detectable parts that constitute the mechanism.
  This test narrows in on two basic epistemic criteria of a mechanistic explanation, namely the context and the parts. It starts by locating the environment within which the phenomenon, and the putative mechanism underlying it, occur. Stipulating the environment should be an easy task because it directly corresponds to the cellular location that researchers choose to study. Furthermore, there are not many choices when it comes to the cell: the intracellular space (i.e. the cytoplasm), the membrane, the extracellular space, and 'sub-environments' relevant to the explanation, for example an organelle such as the mitochondrion inside the cytoplasm.

  The explanation should then stipulate the parts (or entities) of interest. The parts can be detected and measured in the stipulated environment(s). Much of cell biology revolves around molecular parts such as DNAs, RNAs and proteins, which are themselves composed of smaller subparts in the form of nucleic acids, amino acids, etc. There is a reductionist undercurrent in cell biology's reliance on molecular parts for its explanations. The biologist E. O. Wilson explained this as a "search strategy employed to find points of entry into otherwise impenetrably complex systems"

(Wilson, 1998, p. 59). Indeed, as already detailed, much of the AD amyloid explanation relies upon a cast of molecular parts in the form of proteins.

- **TEST #3:** The parts represented in the mechanistic model on which the explanation relies are organised and have some form of interaction.

  This test concerns quantitative associations amongst the parts. Upon detecting the key parts of a mechanism, determining how they are spatiotemporally organised relative to each other and interact amongst each other is a rational next step. In the cell, one could have protein–protein, protein–lipid, protein–DNA and many other types of interactions. Investigating such interactions could be done at a small scale or using what are called 'interactome' approaches, which are aimed at providing a quantifiable snapshot of the interaction of a greater number of molecules. Determining interactions between key parts in a mechanistic model and the resulting explanation may also hint at potential interactions with parts in other (ostensibly unrelated) cellular processes, hence aiding in dealing with the possible masking effect of parallel mechanisms mentioned in **Subsection 2.2**.

  As a lead-up to the next section, consider that if a category of biological principles could be hypothesised that might possibly 'govern' spatiotemporal interactions in the cell, this could have an augmenting role in a PM explanation. More on this later.

- **TEST #4:** The relevant interacting parts described in the explanation exert some form of change on each other via some intermediate means.

  This test stays on the theme of interaction amongst parts but is concerned with the nature of the interactions and the changes they bring about. Put differently, it is essentially about the *consequences* of the associations amongst parts, particularly proteins. Why is this important? A cell's behaviour is ultimately mediated through its plethora of proteins, and hence the physicochemical changes that proteins exert on each other are crucial. It might often turn out that the mere fact of two protein's interaction could be discovered relatively easily, but the biological significance and change(s) brought about by the interaction would take many years to unravel. What also complicates understanding the biological significance of protein–protein interactions is that evolution has created a reality whereby most proteins associate with only a specific subset of other proteins and are relatively inert to interactions with others.[8]

  Now, what could help in establishing the nature of the changes exerted by molecular parts on each other is to analyse the intermediate means of their interactions, which could themselves be other mechanisms (composed of parts, etc.) or things that are not standardly classed as mechanisms, such as physicochemical forces. Proteins are thought to interact with each other via various modes of electrostatic attraction resulting in changes in protein conformation and/or the addition of chemical groups (recall the 'hyperphosphorylation' of tau). In this context, a hypothesised category of biological principles governing macromolecular interactions could inform the test's criteria in a PM explanation.

[8] I am grateful to Gerold Schmitt-Ulms for bringing this point to my attention. See also (Schmitt-Ulms, Mehrabian, Williams, & Ehsani, 2021).

- **TEST #5:** The explanation can accurately account for the sequence of cellular changes leading to detectable variation in the phenomenon.

  Whilst Tests #2 to #4 focus on changes effected on/by individual parts, this test attempts to bridge the parts to the phenomenon by highlighting how a *series* of changes (effected on/by individual parts) may be temporally and causally connected to lead to quantifiable variations in the phenomenon. Indeed, a cornerstone of biological investigations is to study natural or artificial variation/change in a phenomenon as a crucial way of gleaning details about its underlying mechanism.

  To help fulfil this explanatory criterion, one or some cellular part(s) should in principle be manipulated (by the investigator or by nature itself in the case of 'natural experiments'), altering the phenomenon and/or effecting changes on other parts, and paving the way to determining causal relations amongst the parts in the mechanism underlying the phenomenon. Philosophers of science are well-familiar with James Woodward's work on 'interventions' (Woodward, 2016). Whilst interventions in cell biology certainly have to adhere to various methodological standards, the use of 'manipulability' and 'intervention' here are not necessarily committing to the formal constraints of what would count as interventions by Woodward's account.

  As a case in point, suppose that in a cell culture dish of dying neurons, under certain natural conditions, the neurons begin to recover, thus hinting that halting neuronal death in cell culture is possible. Hence, we have a variation in the cell death phenomenon. Let us then assume that an investigator's quantifiable overexpression of a certain protein in dying neurons recapitulates the said variation in the phenomenon, i.e. it halts neuronal death to a certain degree. This now sets the course for researching how, and in what causal sequence of resultant changes, the intervention on (i.e. overexpression of) the protein ultimately leads to the variation in the phenomenon.

These five tests should not be thought of as individually necessary conditions for a productive research programme leading to a successful mechanistic explanation, and, collectively, they are not exhaustive stipulations for a paradigmatic mechanistic explanation. There could surely be further criteria or more exact stipulations for the framework, but the current ones are meant as starting points. Additionally, the tests should not be construed as true-or-false propositions. Even when an explanation fulfils the criteria of a test to a certain degree, the granularity or depth with which those criteria are fulfilled could always be improved upon with refined theoretical work and new empirical findings, and newer questions to investigate could be proposed.

The tests are also not necessarily meant to be hierarchical, but it makes sense for some of them to build on each other. Moreover, a given phenomenon might be explained at least in a rudimentary way even when most of these tests are answered in the negative, that is if the mechanistic model being investigated is at an early stage of development. But in the case of AD, the mechanistic model and explanation have enough detail to be able to engage with the criteria of each test at an appreciable richness.

Below, each of the tests is applied to the amyloid explanation, marking out specific strengths and shortcomings. The point is to systematically ascertain what the explanation might be missing, and how to make it more complete. Furthermore, as introduced above, some of the tests might benefit heuristically from a biological principle that is relevant to their criteria. These opportunities will be highlighted.

- **TEST #1:** *Is the AD **phenomenon** set out as unambiguously as possible?*

  The amyloid explanation has, as one of its endpoints, neuronal damage and cell death (beginning in certain parts of the brain, such as the entorhinal cortex), for which a number of cellular processes leading to Aβ aggregation and NFT formation have been postulated. As far as one could tell, the AD field takes *cellular damage* as the reference phenomenon to investigate the disease. This cellular manifestation of AD is, at the very least, defined in relatively unambiguous terms and communicated as such by investigators in the field. One can therefore tentatively say that, for the most part, the answer to this test is affirmative.

  Having said that, what the amyloid explanation is currently missing is an exact account of how neuronal death at a single-cell level connects with brain-region-specific damage and how that precisely leads to AD's behavioural symptoms. Thus, there is much work to be done, even though this test is being marked as affirmative.

- **TEST #2:** *Does the amyloid explanation set out an **environment** to situate the underlying AD mechanism, and refer to decomposable and detectable **parts** that constitute the mechanism?*

  The amyloid explanation deals within the confines of the environments inside and immediately surrounding single neurons. Furthermore, the explanation rests on a number of key protein players such as APP and tau, which are detectable in experimental settings and whose amino acid sequences are known. Therefore, the answer here is also affirmative.

  Recall, however, that these tests should not be thought of as true-false propositions. For example, what is still missing in the explanation is some account of how, within the densely crowded and highly viscous intracellular environment of the cell, large aggregates of, for instance, tau proteins can even *begin* to form and take up significant intracellular volume. This is not a mere data gap in the explanation, but a *conceptual* gap which the explanation should cover. Thus, here again there is much work to be done, even though the test is being marked as affirmative.

- **TEST #3:** *Are the parts represented in the model on which the amyloid explanation relies **organised** and in some form of **interaction**?*

  In terms of organisation, neurons (which are the 'environment' of the amyloid explanation) are quintessential examples of vastly complex network arrangements of proteins, nucleic acids, lipids and other molecular 'parts'. Moreover, the protein players in the explanation are organised into pathways, as outlined in **Table 1**. Additionally, the chain of direct interactions amongst the parts in either of the (extracellular) Aβ and (intracellular) tau arms of the explanation are clear and known in some depth. However, the interaction of the parts *between* the two arms of the underlying model is much less clear in the explanation (Bloom, 2014; Love, 2001; Rudenko et al., 2019; Tapia-Rojas et al., 2019). Given that both the extracellular and intracellular pathways are implicated in the pathobiology of the disease, there is strong reason based on the existing model to believe there to be important cross-interactions, but there is no indication yet as to their nature. Hence, the tentative answer here for the amyloid explanation is negative. The obvious barrier, and perhaps connection, between the two arms is the *cell membrane*, which spatially separates the internal and external cellular milieus. This possibility will be explored in **Section 5**.

- **TEST #4:** *Do the relevant interacting parts described in the amyloid explanation exert some form of **change** on each other via some **intermediate means**?*

  The question of this test, it may seem, can easily be answered in the affirmative for the cascade of steps that generate the Aβ peptide. For, as detailed in **Section 2**, the very production of the peptide involves changes, e.g. exerted by one part (one of the 'secretase' enzymes) on another part (the APP protein). But what are the intermediate means? The protein structures and 'active sites' (e.g. of the secretase enzymes that cleave APP) are known and have been the subject of many studies (Dehury, Tang, & Kepp, 2019; Seegar et al., 2017); and yet, in terms of exactly how the interaction takes place, we have not progressed much beyond relatively basic appeals to electrostatic interactions and hydrogen bonds. Whilst we can detect a change exerted upon a protein by another, the means with which the change is exerted is nowhere as intelligible as, for example, detecting the parts themselves, their arrangement, etc., and much ground needs to be covered to determine exactly how *chemical* principles, such as electrostatic attraction, are actually operating at the *protein* and *cellular* scales (see e.g. (Matta, 2006; Zhai, Otani, & Ohwada, 2019)). I will therefore mark the answer here as negative. As hinted at earlier, it would be immensely useful if chemical principles of interaction could be transformed into a category of principles that would specifically account for *biological* macromolecular interactions. More on this in **Section 5**.

- **TEST #5:** *Can the amyloid explanation accurately account for the **sequence of cellular changes** that lead to detectable variation in the AD cellular phenomenon?*

  Accounting for detectable variation in the sense of increased protein cleavage, fibrillation and aggregation certainly occupies a central place within the amyloid explanation. Such variations have been painstakingly studied by using, for example, interventions that increase Aβ or tau levels in cultured neurons or in mice and detecting resultant changes in other elements of the presumed underlying mechanism (i.e. parts, interactions, etc.). In addition, certain familial genetic mutations provide *natural* cases of variability of the implicated protein levels. What is still unclear, however, is the certainty and order with which each episode of change can be pinned onto the disease's timeline. As already noted, one of the key open questions in AD research is the extent to which Aβ aggregation is causative of the disease phenomenon or if it is protective (Panza et al., 2019), i.e. whether neuronal cell death starts to happen before or after Aβ aggregation begins in any appreciable manner. The same goes for NFT formation by tau: does it happen before, concomitant with or after plaque formation? Therefore, the sequence of changes (i.e. what happens first, what comes next, etc.) cannot be definitively assigned as of yet, and hence the answer here is negative.

These tests applied in the context of the amyloid explanation can be said to do two things: first, independent of the issue of biological principles, they can systematically help to prioritise which underexplored or missing mechanistic element(s) within the explanation should be investigated and how their discovery could fit into the broader picture of understanding the phenomenon of interest. Second, regardless of the negative/positive assessment of the mechanistic criteria, at least some of the tests could inspire or hint at biological principles that could be discovered, adapted or hypothesised to enrich the overall explanation (i.e.

the PM explanation) and lead to new research questions. Moreover, a given biological principle might feed back into a particular test that inspired its discovery and help resolve the mechanistic gaps that were identified. How principles might achieve these is the topic of the remaining two sections.

5. Two AD-relevant biological principles

I will begin by describing two hypothetical biological principles that could be candidates for inclusion in a PM investigation of AD. These example principles will help with the discussion in **Section 6** where more abstract and general aspects of biological principles will be the focus. I should note that proposing and investigating even one novel biological principle and determining its empirical impact requires a dedicated research programme.

Be that as it may, I will use two of the paradigmatic tests as a springboard here: Test #3, which had to do with the arrangement and interaction of parts, and Test #4, which was centred on interacting parts exerting some form of change on each other via intermediate means. I argued that, respectively, a category of biological principles governing spatiotemporal interactions in the cell and another category accounting for biological macromolecular interactions, could augment our understanding of the criteria picked out in the two tests. To this end, I will propose two principles, provisionally termed the 'principle of cellular synchrony' and the 'principle of generative protein domains', each falling under one of the two categories.

5.1. Principle of cellular synchrony

A few years ago I had suggested that the collective vibrations of 'phospholipids' (a class of lipid molecules) that form the cell membrane may act as a *pacemaker* or timekeeping source for cellular processes at a frequency in the picosecond range (i.e. one-trillionth (10^{-12}) of a second) (Ehsani, 2012). Incidentally, this timekeeping proposal bears similarities to observations that were made by physicists studying biological systems (Adair, 2002; Fröhlich, 1968), indicating that the focus on vibrational behaviour has a clear lineage in biophysics. I should emphasise that this proposal is *not* related to the circadian rhythm (discussed earlier relating to (Bechtel, 2017)), which pertains to the 24-hour timekeeping that happens at an *organism* level and regulates the sleep–wake cycle. Cell membrane vibrations concern an individual cell, and at a time regime that is orders of magnitude faster than one second. Additionally, there is no connection between this proposal and the pacemaker cells of the heart, which are a group of cells that form an electrical conduction system to control the rate of heart muscle contractions in the order of a few seconds.

There are various methods to attempt to validate the link between cell membrane vibrations and timekeeping in the cell. As the evidence stands to-date, the possibility of such a function of the membrane is relatively strong (not discussed further). Nevertheless, building on this possibility, one could posit a *principle of cellular synchrony*, which entails that all processes within a cell (e.g. the activity of proteins, intracellular transport, DNA transcription, etc.) happen in a synchronised manner and are cyclically coordinated based on a subsecond vibrational frequency. This further entails that unsynchronised processes may lead to cell death.

Such a principle, and the notion of a common timing regime, answer to the fact that cellular processes have a mind-boggling level of coordination and interdependence (recall, for example, the earlier discussion of mechanisms 'masking' each other (Illari, 2011)). Moreover, because membranes are a universal feature of cells from all domains of life (Jekely, 2006), this principle would be expected to be operative in any cell. This is my first example of a 'biological principle'.

Why is this a 'biological' rather than a 'chemical' principle? To be sure, the inherent vibrations of every single phospholipid derive from the molecule's *chemistry*. However, when countless phospholipids come together in a membrane, encapsulate the contents of the cell, produce physical vibrations in unison, and that vibrational frequency is transferred across the volume of the cell to affect all the biochemical processes within it, these collectively make the proposed principle a uniquely *biological* one operative at a cellular level. Moreover, the principle could be expanded to entail that in a tissue such as the brain, all adjacent cells, due to direct or indirect membrane contact, might potentially have their cellular processes synchronised as well. This could be an important consequence of the principle because, just as processes in individual cells are intricately coordinated, a collection of neurons and other brain cells also need to be 'in sync' in terms of their synaptic communication, production of action potential, and many other functions. In AD, for example, the loss of synaptic communication (and consequently coordination) between neurons is thought to be a key pathological step as the disease progresses (Edwards, 2019).

Also, in virtue of the above, no *cellular mechanism*[9] could be said to underlie the principle of cellular synchrony. That is, cell membrane vibrations are not the result of a network of interacting proteins and other molecules in the cell. They essentially happen 'automatically' and 'on their own', as an empty shell of just cell membrane would still have the vibrations. However, the vibrations of *each* phospholipid in the membrane could perhaps be explained with a *chemical* or *molecular* mechanistic explanation along with, for example, thermodynamic laws. Thus, even if this principle does admit of further explanation, the explanation is unlikely to be a purely mechanistic one — appealing as it presumably will to thermodynamic laws — and any mechanistic component of the explanation is likely to be at the chemical rather than the biological level.

Now, on the basis of the temporal synchrony that this principle could entail across cellular processes, we might consider it as a member of potential principles of *spatiotemporal interactions*. Going back to Test #3 on how the membrane could bridge the extracellular and intracellular facets of the AD amyloid explanation, a hypothesis that the principle of synchrony brings forth is that both the Aβ peptide and hyperphosphorylated tau might interfere in tandem with the vibrations of the membrane. We know, for example, that extracellular Aβ oligomers "destabilize the […] membrane's structure, induce a generalized increase in membrane permeability, and insert themselves into the membrane to form cation-conducting pores" (Wang, Tan, Liu, & Yu, 2016, p. 1914). It has also been reported that the tau protein forms complexes with phospholipids (Ait-Bouziad et al., 2017). And, Aβ might in fact directly interact with and facilitate the fibrilization of tau (Vasconcelos et al., 2016).

The interplay of unsynchronised cellular processes and AD pathology, mediated by the membrane and both Aβ and tau, would be a possible new path of investigation. This could lead to novel insights on (i)

[9] Here I intend for 'cellular' mechanisms to be distinct from physical or chemical mechanisms.

how the main parts and pathways of the AD mechanistic model could be explained to be connected, and (ii) how the pathways interact and influence each other. This is one scenario in which a new biological principle could *heuristically apply to the elements of a mechanistic explanation*, in this case showing one way of allowing the AD explanation to satisfy the conditions of Test #3. Furthermore, the principle can *independently (and without necessarily invoking the mechanistic elements)* raise the explanatory possibility that AD could, for instance, signal an overall unsynchronised timing amongst cellular processes in a given region of the brain. If such a line of investigation were to be pursued, *quantifications* of the unsynchronised timing of cellular processes might also be studied: in light of the loss of synaptic communication and coordination, an AD researcher could, for instance, investigate how long certain cellular processes might be 'off' relative to each other, and if this timing difference would increase as the disease progresses.

5.2. Principle of generative protein domains

A second hypothetical biological principle could relate to the geometry and three-dimensional conformation of proteins such as Aβ and tau. In this subsection, I would like to specifically focus on tau, which is taken to be an 'intrinsically disordered' protein (Sabbagh & Dickey, 2016).

Since the 1950s, protein structure has been defined in terms of *domains* or segments that may be of (roughly) three flavours: 'alpha-helical', 'beta-sheet' and 'disordered'. Approximately a third of human proteins are thought to contain disordered segments, and these segments contribute significantly to protein–protein interactions and therefore to protein function (Kim & Chung, 2020). A visual representation of how ordered versus disordered protein domains can be thought of is presented in **Figure 2** on the next page. A persistent challenge, however, has been that there is no overarching explanatory theory of what a disordered domain is and how it could systematically be conceptualised.

But what if disordered proteins are not really 'disordered' after all? Commentators on the physics of 'order' and 'chaos', or 'patterned' and 'patternless', do not usually entertain a *third* possibility, one which would in principle be neither ordered nor disordered (Crutchfield, 2012). Truly envisioning what a third possibility could be may perhaps be beyond human cognitive capacity, but there have been attempts toward such an aim (see e.g. (Clouser et al., 2019; Sormanni et al., 2017)).

Within this context, one could propose a *principle of generative protein domains*, stating that so-called disordered protein domains might possibly be 'generative' domains in that they can assume *multiple* precise structures that are appropriate to their immediate cellular environment and interacting partners. The possible structures that a disordered domain could assume might be dependent on the complex balance of a number of factors, including: (i) the *chemical* interactions (electrostatic and non-electrostatic) within the protein domain itself (amongst its different amino acids) and between different domains of the same or another protein, (ii) the *geometrical* constraints of the protein (e.g. some structural configurations might simply be geometrically impossible for the protein macromolecule), and (iii) the *physicochemical* properties of the cellular fluid/medium surrounding the protein (its composition, viscosity, etc.) and those properties' effects on the domain's structure.

I should note that the factors stipulated for this principle could potentially be operative on any protein domain, regardless of its 'orderedness'. However, it is reasonable to assume that for ordered domains (alpha-helical and beta-sheet), the influence of the first factor (i.e. chemical interactions amongst a domain's amino acids) might be so overwhelming so as to outweigh the contributions of the other two factors. This is due to the prevailing notion that the single structure of ordered domains can be entirely explained by chemical properties at the protein sequence level (Söding, 2017), and that these domains assume the same 'constant' structure in most cellular contexts. As a further note, the principle of generative protein domains can be assigned to the biological macromolecular interactions category, and, since it is wholly concerned with proteins in cells, it is unambiguously a *biological* principle operative at a cellular level.

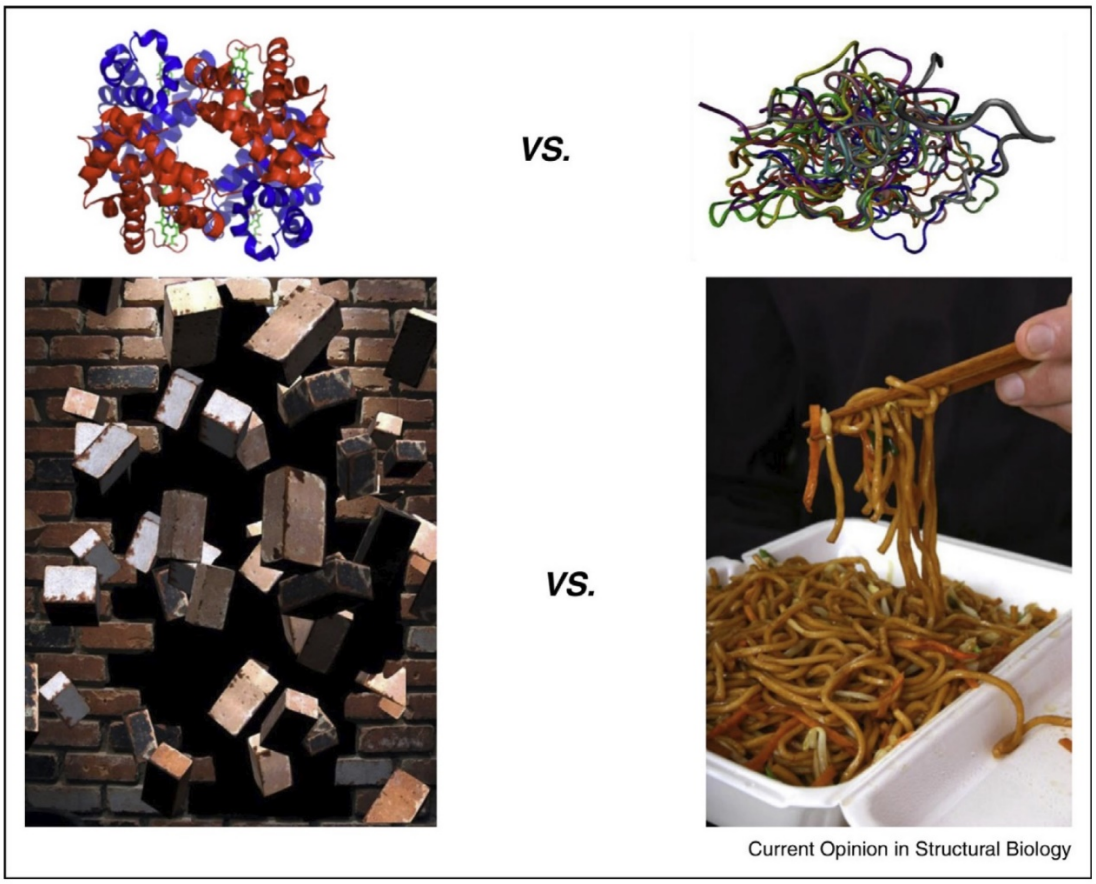

Figure 2. Representations of ordered and disordered proteins. This figure from (Uversky, 2017) depicts a multidomain ordered protein (an alpha-helical protein) on the top left and a disordered protein on the top right. The structural and conformational rigidity of the *domains* within these two types of protein structure is compared to bricks vs. pasta, respectively. The question, however, is that is a disordered-domain-containing protein — such as the tau protein (which is actually entirely disordered) — really without any particularly constant shape (like a string of pasta), or is it perhaps 'structured' in some other way? Reproduced with permission from Elsevier.

Now, could this principle, similar to the principle of cellular synchrony, also be counted as ontologically distinct from cellular mechanistic explanations? I argue that it can, given that it fundamentally concerns the internal and external *chemistry* of proteins, and as such it might only be possible to posit *chemical* mechanisms underwriting it, if at all. Consider that we are not dealing with a small molecule interacting with its surrounding medium in the cytoplasm or the geometrical constraints of a few atoms. A protein is a *macro*molecule with

potentially many properties that cannot entirely be explained by chemical mechanisms and thermodynamic laws (Jiang, Teufel, Jackson, & Wilke, 2018).

Moving forward, and taking the principle of generative protein domains as *governing* the structure of disordered protein domains, one could posit that in *certain cellular milieus*, the tau protein may indeed have a precise and constant structure (i.e. a constrained structure without any appreciable variance), albeit not alpha-helical or beta-sheet ordered. One implication is that the tau protein sequence might be able to 'generate' different three-dimensional structures depending on its hyperphosphorylation status (a geometric and chemical change), and/or whether it is in the intracellular fluid or interacting with the membrane (a medium change). This could actually also go some way in accounting for the *multiple roles* that a disordered protein could have in the cell (see e.g. (Olivieri et al., 2020)). The generative protein domains principle could thus help to *explain facets of the behaviour of tau in AD neurons independent of the amyloid explanation*.

At the same time, however, the potential implications of this principle could inform the problem highlighted in Test #4. There, a problem was that the current conception of how two proteins interact is tenuous: "in terms of exactly how the interaction takes place, we have not progressed much beyond relatively basic appeals to electrostatic interactions and hydrogen bonds" (**Section 4**). Obtaining new insights on the dynamics of tau's structure could lead to clues about its interaction with other 'ordered' and 'disordered' proteins. The principle might also eventually lead to *quantifications* of the extent to which the interaction between such proteins is dictated by electrostatic forces and hydrogen bonds, and how much — as suggested by the principle's stipulations — is shaped by the medium surrounding the proteins and the mutual geometrical constraints of their interacting domains. We might thus be able to quantify the *strength* of tau's interaction with its key interacting protein partners in the AD amyloid explanation.

The principles of cellular synchrony and generative protein domains are just two examples of potential biological principles that could be hypothesised based on the problems at hand. For each of these principles, one could ask what *parameters* (i.e. context-dependent and changeable factors similar to a temperature variable in a thermodynamic law) could fine-tune the principle. In the case of the cellular synchrony principle, the 'rigidity' of the cell membrane could be a possible parameter, whereby the more rigid the membrane, the lower its vibrational frequency might be expected to be. The rigidity could be dictated by, for instance, how tightly packed the membrane is of phospholipids. In the case of the generative domains principle, potential parameters are indeed stipulated in the principle itself, such as the composition or the viscosity of the medium surrounding the protein. All such parameters need to be validated experimentally.

In all, the examples in this section were meant to demonstrate how the discovery of biological principles could go hand-in-hand with the elucidation of the mechanistic elements of a mechanistic explanation brought about by the paradigmatic tests. Moreover, the two hypothesised principles showed how biological principles might have non-mechanistic along with mechanistic effects on the overall explanation: they might do independent explanatory work, and they might also act as heuristics for making mechanistic explanations more 'complete'. The next section abstracts away from the specifics of the examples in this section to draw more general claims about the place of biological principles in a PM explanation.

6. Biological principles and mechanistic explanations

The examples of the principles of cellular synchrony and generative protein domains can help crystallise some general properties about biological principles. Starting with their non-mechanistic effects on the overall explanation, one might initially ask how biological principles would fit into the broader notion of scientific laws. Indeed, in the special sciences (basically any natural science other than fundamental physics), it is hard to come by generalisations that can act as almost universal laws that, as noted previously, are characterised by "universality, inviolable necessity, or unrestricted scope" (Craver & Tabery, 2019).[10]

As such, philosophers of science usually discuss patterns and generalisations in fields like biology under the framework of *ceteris paribus* laws, i.e. laws that hold when *other things are (held) equal*. These laws are essentially generalisations that admit of various context-dependent exceptions, and therefore can be thought of as a category of non-exceptionless generalisations (Fenton-Glynn, 2016). The discussion of *ceteris paribus* laws traces its history to the economic sciences but has been applied and critiqued in all branches of the special sciences. For reasons that will be discussed elsewhere, in this paper the generalisations concerning cell biology which I call 'principles' approximate what some philosophers of science mean by *ceteris paribus* laws or 'invariant generalisations'.[11]

Some further points also need to be made about the word 'biological' in 'biological principles'. First, principles or laws from the broader *life sciences* could be applied to cell biology. Foremost amongst these are evolutionary principles (Linquist et al., 2016), which generally apply at a *cell population* level. Indeed, many forms of generalisations that cell biologists may know as 'principles' or 'laws' fall within the purview of observational patterns or conjectures in (i) evolution (e.g. concerning genetic plasticity (Hannan, 2018)), (ii) zoology (e.g. concerning tissue patterning (Barkai & Shilo, 2020)), (iii) ecology or (iv) biogeography, and can often have mathematical form. Second, a motivation for the PM framework is to argue that biological principles for cellular phenomena should be recognised as being distinct from — but connected to — thermodynamic, evolutionary, ecological and other related laws and principles.[12]

---

[10] An example is Newton's law of universal gravitation. Even then, however, one needs to be cognisant of the effect of electrical force between charged bodies (see (Elgin, 2017, p. 25)).

[11] Specifically, as I will argue in subsequent work, a 'principle' can be thought of as a generalisation of some perceived regularity in the natural system being studied that is not yet at the level of a law claim. If it eventually turns out to be an accidental generalisation, for example, based on insufficient evidence at the time it was posited, then it could get discarded or modified. One or more principles, over time and with mounting evidence, could be formalised into a law claim. See also (Berryman, 2003) for a related analysis concerning laws in population ecology.

[12] Arguments regarding the distinctiveness of biological principles could be varied and draw from many sources. As a case in point, Marc Lange, in a chapter on "what would natural laws in the life sciences be", writes that "a biological generalization can possess a distinctive variety of necessity – can be a biological law. Associated with this distinctive necessity is a range of invariance under counterfactual antecedents that is broader in some respects than the range of invariance exhibited by the fundamental physical laws" (Lange, 2013, p. 83). Beyond broadness, the distinctiveness of biological principles can also stem from the uniqueness of *cellular* phenomena compared to physical and chemical phenomena. Of note, the discussion of biological principles could just as well cover *organism*-level phenomena, but an in-depth explanation of such phenomena (bearing in mind that an organism is made up of a vast number of different cells) would necessarily be much more complex. As such, here I have restricted the scope of biological principles only to the phenomena of the cell, which is biology's basic unit of study.

Having these distinctions at hand, how could biological principles deliver independent explanatory work? To begin with, consider, for example, that the principle of cellular synchrony could potentially account for the collective non-synchronised behaviour of a group of adjacent neurons in AD patients without referencing any particular mechanistic explanation, such as the amyloid explanation. The principle could stipulate a common reaction timing regime for some number of cells and, if an investigator so chooses, 'explain' AD cellular phenomena completely non-mechanistically. This line of study could also facilitate the prediction and 'expectability' (Deulofeu & Suárez, 2018) of what might happen in the brain as the disease progresses.[13]

This brings to mind the way in which thermodynamic laws allow a chemist to feel confident, to a certain extent, about what to expect before a reaction is initiated. This does not mean that thermodynamic laws would necessarily lead to the exact prediction of a reaction's outcome (although that is certainly an ultimate goal), but rather whether to simply expect, for example, the reaction to be endothermic or not. Importantly, the predictive power facilitated by biological principles might additionally make the explanation more quantitative, as was described for the two examples: the principle of cellular timekeeping could lead to measurements in cellular process time differences, and the principle on disordered protein domains could lead to better quantifications of the strength of protein–protein interactions.

Moving next to the mechanistic explanatory effects of biological principles, recall from **Section 4** that at least some of the mechanistic criteria picked out by the paradigmatic tests could in theory be governed by one or more hypothesised biological principles. The principles could help general information to bear on the mechanistic explanation and thus be utilised to (i) understand its various facets, such as the nature of the parts or their interactions (as the generative protein domains principle did for protein structure), and (ii) move toward the resolution of the potentially missing elements of the mechanistic explanation (as the cellular synchrony principle did to bridge the extracellular and intracellular facets of the amyloid explanation).

The mechanistic explanatory effects of principles, I would argue, might complement and fit into at least a segment of current NM investigations, such as the principles of organisation (or design principles) that were alluded to earlier (Bechtel, 2017). However, to what extent this generalises across NM-related investigations is a matter of debate. As quoted in **Section 3**, Cartwright and colleagues are of the opinion that a mechanism's 'regular behaviour' "is what it takes for some set of principles that govern the features of M's parts in their arrangement in M all to be instanced together" (Cartwright et al., 2020, p. 18). On the other hand, Beate Krickel argues that "the new mechanists could in principle accept that […] the interactions between mechanistic components are governed by laws (which they usually do not) but that it does not follow that laws or expectability adds any explanatory power to a mechanistic explanation" (Krickel, 2020, p. 8). This paper's arguments, however, suggest that biological principles have the potential to increase the explanatory power of mechanistic explanations of cellular phenomena in a heuristic manner. Furthermore, the principles themselves might independently explain certain facets of the target phenomenon, thereby increasing the explanatory power of the *overall* explanation (which I have called the PM explanation).

[13] There is pertinent discussion to be had about the independent explanatory work of biological principles and the literature on 'grounding' (e.g. (Trogdon, 2018)), but one which I shall postpone to a separate analysis.

To recap, the goal of the PM framework, after having outlined what a paradigmatic mechanistic explanation could be, is to discover and apply biological principles that are operative in a biological cell. The principles could have independent explanatory power, aiding in prediction and expectability. Biological principles could also make various components of the mechanistic explanation more complete, thereby also increasing the depth of the mechanistic share of the overall explanation.

7. Conclusions

In this paper, following a review of the significant problems faced by the Alzheimer's research field, I argued that various strands of clinical evidence point to the conclusion that the existing and widely-accepted amyloid cascade mechanistic explanation appears to be fundamentally incomplete. In this context, I then proposed that a theoretical framework called 'principled mechanism' (PM) has the potential to inspire new sets of empirical questions and novel avenues of investigation that can take a given mechanistic explanation at any stage of development as its starting point. PM is a two-pronged framework: on the one hand, using a short series of 'tests', it systematically compares different components of the mechanistic explanation against a paradigmatic set of criteria, and hints at various ways of making the mechanistic explanation more complete. On the other hand, it makes an appeal for the discovery and application of 'biological principles' relevant to the phenomenon being explained. Next, I detailed two hypothetical biological principles, one having to do with 'time' at the cellular level and the other concerning the structure of 'disordered' protein domains, and how they could each inform and improve different aspects of the amyloid explanation of AD. Using these examples, I argued that biological principles are 'principles' since they approximate what some philosophers mean by *ceteris paribus* laws or 'invariant generalisations', and 'biological' because they are operative at and specific to the level of a biological cell. Such principles could aid in several ways, such as augmenting different facets of the mechanistic explanation itself but also allowing further independent nomological explanation of the phenomenon. Whilst this strategy can be complementary to certain New Mechanist approaches, an important distinction of the PM framework is its equal attention to the explanatory utility of biological principles.

Overall, PM can help to move cell biological investigations from what might be called a generally 'mechanistic-descriptive' state (the status quo) to a 'mechanistic-nomological' paradigm, entailing theoretic biological principles alongside mechanistic accounts. If the purpose of an AD research programme, in our case, is to arrive at a deeper form of biological explanation toward effective treatments, and in doing so introduce the field to hitherto uninvestigated paths of enquiry, then the PM approach can be useful in making the current AD amyloid model and resultant explanation much more adequate-for-purpose (see (Parker, 2020)). Moreover, the PM framework could just as easily be applied to other biomedical domains, such as mechanistic explanations in cancer research, diabetes or cardiovascular disease, but also explanations for physiological cellular phenomena.

Ultimately, and notwithstanding the issues of the importance or impact of PM, a broader goal of the case study is to show that what is perceived by some to be a fragmentation between analytic philosophy and biomedicine is artificial and unhelpful, and that progress can be made when philosophical approaches are directed at open and neglected problems concerning the biological cell.

Acknowledgements

I would like to convey my many thanks to Luke Fenton-Glynn (UCL Philosophy) and Phyllis Illari (UCL Science and Technology Studies) for their careful reading of and very helpful comments on the draft versions of this manuscript. This work and the motivation to study the theoretical basis of Alzheimer's disease are dedicated to the memory of Eshrat Khazaeinejad.

References

Adair, R. K. (2002). Vibrational resonances in biological systems at microwave frequencies. *Biophys J, 82*(3), 1147–1152.

Aisen, P. (2019). Comment on "End of the BACE Inhibitors? Elenbecestat Trials Halted Amid Safety Concerns" by J. Shugart. *Alzforum*. Retrieved from alzforum.org/news/research-news/end-bace-inhibitors-elenbecestat-trials-halted-amid-safety-concerns#comment-32966

Ait-Bouziad, N., Lv, G., Mahul-Mellier, A. L., Xiao, S., Zorludemir, G., Eliezer, D., . . . Lashuel, H. A. (2017). Discovery and characterization of stable and toxic Tau/phospholipid oligomeric complexes. *Nat Commun, 8*(1), 1678.

Alleva, K., Diez, J., & Federico, L. (2017). Models, theory structure and mechanisms in biochemistry: The case of allosterism. *Stud Hist Philos Biol Biomed Sci, 63*, 1–14.

Alzheimer, A., Stelzmann, R. A., Schnitzlein, H. N., & Murtagh, F. R. (1995). An English translation of Alzheimer's 1907 paper, "Uber eine eigenartige Erkankung der Hirnrinde". *Clin Anat, 8*(6), 429–431.

Arendt, T., Stieler, J., Strijkstra, A. M., Hut, R. A., Rudiger, J., Van der Zee, E. A., . . . Hartig, W. (2003). Reversible paired helical filament-like phosphorylation of tau is an adaptive process associated with neuronal plasticity in hibernating animals. *J Neurosci, 23*(18), 6972–6981.

Barkai, N., & Shilo, B. Z. (2020). Reconstituting tissue patterning. *Science, 370*(6514), 292–293.

Bechtel, W. (2011). Mechanism and Biological Explanation. *Philosophy of Science, 78*(4), 533–557.

Bechtel, W. (2017). Explaining features of fine-grained phenomena using abstract analyses of phenomena and mechanisms: two examples from chronobiology. *Synthese*(S.I. : Abstraction and Idealization in Scientific Modelling), 1–23.

Bechtel, W., & Abrahamsen, A. (2005). Explanation: a mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences, 36*(2), 421–441.

Bechtel, W., & Richardson, R. C. (2010). *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*: MIT Press.

Berryman, A. A. (2003). On principles, laws and theory in population ecology. *Oikos, 103*(3), 695–701.

Beyreuther, K., & Masters, C. L. (1991). Amyloid precursor protein (APP) and beta A4 amyloid in the etiology of Alzheimer's disease: precursor-product relationships in the derangement of neuronal function. *Brain Pathol, 1*(4), 241–251.

Blackburn, S. (2016). *The Oxford Dictionary of Philosophy* (3rd ed.): Oxford University Press.

Bloom, G. S. (2014). Amyloid-beta and tau: the trigger and bullet in Alzheimer disease pathogenesis. *JAMA Neurol, 71*(4), 505–508.

Bogen, J. (2020). Theory and Observation in Science. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.

Brini, E., Simmerling, C., & Dill, K. (2020). Protein storytelling through physics. *Science, 370*(6520).

Cairns, D. M., Rouleau, N., Parker, R. N., Walsh, K. G., Gehrke, L., & Kaplan, D. L. (2020). A 3D human brain-like tissue model of herpes-induced Alzheimer's disease. *Sci Adv, 6*(19), eaay8828.

Cartwright, N., Pemberton, J., & Wieten, S. (2020). Mechanisms, laws and explanation. *European Journal for Philosophy of Science, 10*(25).

Clouser, A. F., Baughman, H. E., Basanta, B., Guttman, M., Nath, A., & Klevit, R. E. (2019). Interplay of disordered and ordered regions of a human small heat shock protein yields an ensemble of 'quasi-ordered' states. *Elife, 8*.

Craver, C. F., & Kaiser, M. I. (2013). Mechanisms and Laws: Clarifying the Debate. In H.-K. Chao, S.-T. Chen & R. L. Millstein (Eds.), *Mechanism and Causality in Biology and Economics* (pp. 125–145): Springer.

Craver, C. F., & Tabery, J. (2019). Mechanisms in Science. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.

Crutchfield, J. P. (2012). Between order and chaos. *Nature Physics, 8*, 17–24.

Dear, A. J., Michaels, T. C. T., Meisl, G., Klenerman, D., Wu, S., Perrett, S., . . . Knowles, T. P. J. (2020). Kinetic diversity of amyloid oligomers. *Proc Natl Acad Sci U S A, 117*(22), 12087–12094.

Dehury, B., Tang, N., & Kepp, K. P. (2019). Molecular dynamics of C99-bound gamma-secretase reveal two binding modes with distinct compactness, stability, and active-site retention: implications for Abeta production. *Biochem J, 476*(7), 1173–1189.

Dekkers, W., & Rikkert, M. O. (2006). What is a genetic cause? The example of Alzheimer's Disease. *Med Health Care Philos, 9*(3), 273–284.

Deulofeu, R., & Suárez, J. (2018). When Mechanisms Are Not Enough: The Origin of Eukaryotes and Scientific Explanation. In A. Christian, D. Hommen, N. Retzlaff & G. Schurz (Eds.), *Philosophy of Science: Between the Natural Sciences, the Social Sciences, and the Humanities* (pp. 95–115): Springer.
Díez, J. (2014). Scientific w-Explanation as Ampliative, Specialized Embedding: A Neo-Hempelian Account. *Erkenntnis, 79*, 1413–1443.
Dijksterhuis, E. J. (1961). *The Mechanization of the World Picture* (C. Dikshoorn, Trans.): Oxford University Press.
Editors. (2016). Dementia across the Lifespan and around the Globe-Pathophysiology, Prevention, Treatment, and Societal Impact: A Call for Papers. *PLoS Med, 13*(8), e1002117.
Edwards, F. A. (2019). A Unifying Hypothesis for Alzheimer's Disease: From Plaques to Neurodegeneration. *Trends Neurosci, 42*(5), 310-322.
Ehsani, S. (2012). Time in the cell: a plausible role for the plasma membrane. *arXiv, 1210.0168*.
Elgin, C. Z. (2017). *True Enough*: MIT Press.
Feinman, R. D., & Fine, E. J. (2004). "A calorie is a calorie" violates the second law of thermodynamics. *Nutr J, 3*, 9.
Fenton-Glynn, L. (2016). *Ceteris Paribus* Laws and *Minutis Rectis* Laws. *Philos Phenomenol Res, 93*(2), 274–305.
Franklin-Hall, L. R. (2016). New Mechanistic Explanation and the Need for Explanatory Constraints. In K. Aizawa & C. Gillett (Eds.), *Scientific Composition and Metaphysical Ground* (pp. 41–74): Palgrave Macmillan.
Frigg, R., & Nguyen, J. (2016). The Fiction View of Models Reloaded. *The Monist, 99*(3), 225–242.
Fröhlich, H. (1968). Long-range coherence and energy storage in biological systems. *Int J Quantum Chem, 2*(5), 641–649.
Glennan, S. (2017). *The New Mechanical Philosophy*: Oxford University Press.
Glennan, S., & Illari, P. (2017). Varieties of mechanisms. In S. Glennan & P. Illari (Eds.), *The Routledge Handbook of Mechanisms and Mechanical Philosophy* (pp. 91–103): Routledge.
Green, S. (2015). Revisiting generality in biology: systems biology and the quest for design principles. *Biol Philos, 30*, 629–652.
Hammond, M., & Newton, T. (2020). We need to radically rethink our approach to Alzheimer's research. *Wired UK*. Retrieved from wired.co.uk/article/alzheimers-research
Hannan, A. J. (2018). Tandem repeats mediating genetic plasticity in health and disease. *Nat Rev Genet, 19*(5), 286–298.
Hardy, J., & Allsop, D. (1991). Amyloid deposition as the central event in the aetiology of Alzheimer's disease. *Trends Pharmacol Sci, 12*(10), 383–388.
Hempel, C. G. (1965). Aspects of Scientific Explanation *Aspects of Scientific Explanation and other Essays in the Philosophy of Science* (pp. 331–496): Free Press.
Henderson, N. C., Rieder, F., & Wynn, T. A. (2020). Fibrosis: from mechanisms to medicines. *Nature, 587*(7835), 555–566.
Illari, P. (2011). Mechanistic Evidence: Disambiguating the Russo-Williamson Thesis. *International Studies in the Philosophy of Science, 25*(2), 139–157.
Ishii, M., & Iadecola, C. (2020). Risk factor for Alzheimer's disease breaks the blood-brain barrier. *Nature, 581*(7806), 31–32.
Jekely, G. (2006). Did the last common ancestor have a biological membrane? *Biol Direct, 1*, 35.
Jiang, Q., Teufel, A. I., Jackson, E. L., & Wilke, C. O. (2018). Beyond Thermodynamic Constraints: Evolutionary Sampling Generates Realistic Protein Sequence Variation. *Genetics, 208*(4), 1387-1395.
Kim, J. Y., & Chung, H. S. (2020). Disordered proteins follow diverse transition paths as they fold and bind to a partner. *Science, 368*(6496), 1253–1257.
Kontos, P. C. (2004). Ethnographic reflections on selfhood, embodiment and Alzheimer's disease. *Ageing & Society, 24*, 829–849.
Krickel, B. (2020). Reply to Cartwright, Pemberton, Wieten: "mechanisms, laws and explanation". *European Journal for Philosophy of Science, 10*(43), 1–9.
Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*: The University of Chicago Press
Kundu, K., Darden, L., & Moult, J. (2020). MecCog: A knowledge representation framework for genetic disease mechanism. *bioRxiv, 282012*.
Lange, M. (2013). What Would Natural Laws in the Life Sciences Be? In K. Kampourakis (Ed.), *The Philosophy of Biology: A Companion for Educators* (pp. 67–85): Springer.
Leuridan, B. (2010). Can Mechanisms Really Replace Laws of Nature? *Philosophy of Science, 77*(3), 317–340.
Linquist, S., Gregory, T. R., Elliott, T. A., Saylor, B., Kremer, S. C., & Cottenie, K. (2016). Yes! There are resilient generalizations (or "laws") in ecology. *Q Rev Biol, 91*(2), 119–131.
Liu, P. P., Xie, Y., Meng, X. Y., & Kang, J. S. (2019). History and progress of hypotheses and clinical trials for Alzheimer's disease. *Signal Transduct Target Ther, 4*, 29.
Llored, J.-P. (2011). The Role and the Status of Thermodynamics in Quantum Chemistry Calculations. In J. C. Moreno-Piraján (Ed.), *Thermodynamics: Interaction Studies – Solids, Liquids and Gases* (pp. 469–490): IntechOpen.
Lombrozo, T. (2011). The Instrumental Value of Explanations. *Philosophy Compass, 6*(8), 539–551.
Love, R. (2001). Untangling the relation between beta-amyloid and tau. *Lancet, 358*(9282), 645.
Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about Mechanisms. *Philosophy of Science, 67*(1), 1–25.
Marder, E. (2020). Words without meaning. *Elife, 9*.
Masters, C. L. (2019). Comment on "Electrode Detects Aβ Aggregates in Alzheimer's Plasma" by G. D. Zakaib. *Alzforum*. Retrieved from alzforum.org/news/research-news/electrode-detects-av-aggregates-alzheimers-plasma#comment-31421

Matta, C. F. (2006). Hydrogen-hydrogen bonding: The non-electrostatic limit of closed-shell interaction between two hydrogen atoms. A critical review. In S. J. Grabowski (Ed.), *Hydrogen Bonding - New Insights* (pp. 337–375): Springer.
Moir, R. D., Lathe, R., & Tanzi, R. E. (2018). The antimicrobial protection hypothesis of Alzheimer's disease. *Alzheimers Dement, 14*(12), 1602–1614.
Newton-Smith, W. H. (1981). *The Rationality of Science*: Routledge & Kegan Paul.
Nordenfelt, L. (2006). Commentary on Wim Dekkers's and Marcel Olde Rikkert's: "What is a genetic disease? The example of Alzheimer's Disease", and Stephen Tyreman's: "Causes of illness in clinical practice: a conceptual exploration". *Med Health Care Philos, 9*(3), 317–319.
NTNU. (2020). Norway's Nobel laureates take up the fight against Alzheimer's. *EurekAlert*. Retrieved from eurekalert.org/pub_releases/2020-03/nuos-nnl030420.php
Olivieri, C., Wang, Y., Li, G. C., V, S. M., Kim, J., Stultz, B. R., . . . Veglia, G. (2020). Multi-state recognition pathway of the intrinsically disordered protein kinase inhibitor by protein kinase A. *Elife, 9*.
Panza, F., Lozupone, M., Logroscino, G., & Imbimbo, B. P. (2019). A critical appraisal of amyloid-beta-targeting therapies for Alzheimer disease. *Nat Rev Neurol, 15*(2), 73–88.
Park, L., Hochrainer, K., Hattori, Y., Ahn, S. J., Anfray, A., Wang, G., . . . Iadecola, C. (2020). Tau induces PSD95-neuronal NOS uncoupling and neurovascular dysfunction independent of neurodegeneration. *Nat Neurosci, 23*(9), 1079–1089.
Parker, W. S. (2020). Model Evaluation: An Adequacy-for-Purpose View. *Philosophy of Science, 87*(3), 457–477.
Pastore, A., Raimondi, F., Rajendran, L., & Temussi, P. A. (2020). Why does the Abeta peptide of Alzheimer share structural similarity with antimicrobial peptides? *Commun Biol, 3*(1), 135.
Pearl, J., & Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*: Basic Books.
Petsko, G. A. (2018). Comment on "Crystal Structure of Aβ and Proposed Receptor Solved" by G. D. Zakaib. *Alzforum*. Retrieved from alzforum.org/news/research-news/crystal-structure-av-and-proposed-receptor-solved#comment-29171
Richardson, J. O., Perez, C., Lobsiger, S., Reid, A. A., Temelso, B., Shields, G. C., . . . Althorpe, S. C. (2016). Concerted hydrogen-bond breaking by quantum tunneling in the water hexamer prism. *Science, 351*(6279), 1310–1313.
Rudenko, L. K., Wallrabe, H., Periasamy, A., Siller, K. H., Svindrych, Z., Seward, M. E., . . . Bloom, G. S. (2019). Intraneuronal Tau Misfolding Induced by Extracellular Amyloid-beta Oligomers. *J Alzheimers Dis, 71*(4), 1125–1138.
Sabbagh, J. J., & Dickey, C. A. (2016). The Metamorphic Nature of the Tau Protein: Dynamic Flexibility Comes at a Cost. *Front Neurosci, 10*, 3.
Schmitt-Ulms, G., Mehrabian, M., Williams, D., & Ehsani, S. (2021). The IDIP framework for assessing protein function and its application to the prion protein. *Biological Reviews*.
Seegar, T. C. M., Killingsworth, L. B., Saha, N., Meyer, P. A., Patra, D., Zimmerman, B., . . . Blacklow, S. C. (2017). Structural Basis for Regulated Proteolysis by the alpha-Secretase ADAM10. *Cell, 171*(7), 1638–1648 e1637.
Selkoe, D. J. (1991). The molecular pathology of Alzheimer's disease. *Neuron, 6*(4), 487–498.
Shugart, J. (2020). In APP Knock-In Mice, Spatial Mapping Circuitry Crumbles. *Alzforum*. Retrieved from alzforum.org/news/research-news/app-knock-mice-spatial-mapping-circuitry-crumbles-0
Söding, J. (2017). Big-data approaches to protein structure prediction. *Science, 355*(6322), 248–249.
Sormanni, P., Piovesan, D., Heller, G. T., Bonomi, M., Kukic, P., Camilloni, C., . . . Vendruscolo, M. (2017). Simultaneous quantification of protein order and disorder. *Nat Chem Biol, 13*(4), 339–342.
Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Dev Sci, 10*(1), 89–96.
Tapia-Rojas, C., Cabezas-Opazo, F., Deaton, C. A., Vergara, E. H., Johnson, G. V. W., & Quintanilla, R. A. (2019). It's all about tau. *Prog Neurobiol, 175*, 54–76.
Theurer, K. L. (2018). 'Looking Up' and 'Looking Down': On the Dual Character of Mechanistic Explanations. *J Gen Philos Sci, 49*, 371–392.
Trogdon, K. (2018). Grounding-mechanical explanation. *Philosophical Studies, 175*, 1289–1309.
Uversky, V. N. (2017). Intrinsically disordered proteins in overcrowded milieu: Membrane-less organelles, phase separation, and intrinsic disorder. *Curr Opin Struct Biol, 44*, 18–30.
Vasconcelos, B., Stancu, I. C., Buist, A., Bird, M., Wang, P., Vanoosthuyse, A., . . . Dewachter, I. (2016). Heterotypic seeding of Tau fibrillization by pre-aggregated Abeta provides potent seeds for prion-like seeding and propagation of Tau-pathology in vivo. *Acta Neuropathol, 131*(4), 549–569.
Wang, Z. X., Tan, L., Liu, J., & Yu, J. T. (2016). The Essential Role of Soluble Abeta Oligomers in Alzheimer's Disease. *Mol Neurobiol, 53*(3), 1905–1924.
Wilson, E. O. (1998). *Consilience: The Unity of Knowledge*: Alfred A. Knopf.
Wimsatt, W. C. (2017). Foreword. In S. Glennan & P. Illari (Eds.), *The Routledge Handbook of Mechanisms and Mechanical Philosophy* (pp. xiv–xvi): Routledge.
Woodward, J. (2016). Causation and Manipulability. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.
Yadav, A., Vidal, M., & Luck, K. (2020). Precision medicine - networks to the rescue. *Curr Opin Biotechnol, 63*, 177–189.
Yamazaki, Y., Zhao, N., Caulfield, T. R., Liu, C. C., & Bu, G. (2019). Apolipoprotein E and Alzheimer disease: pathobiology and targeting strategies. *Nat Rev Neurol, 15*(9), 501–518.
Zhai, L., Otani, Y., & Ohwada, T. (2019). Uncovering the Networks of Topological Neighborhoods in beta-Strand and Amyloid beta-Sheet Structures. *Sci Rep, 9*(1), 10737.