A transcriptional signature of Alzheimer's disease is associated with a metastable subproteome at risk for aggregation

Prajwal Ciryam[a,b,1], Rishika Kundra[a,1], Rosie Freer[a], Richard I. Morimoto[b], Christopher M. Dobson[a], and Michele Vendruscolo[a,2]

[a]Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom; and [b]Department of Molecular Biosciences, Northwestern University, Evanston, IL 60208

Edited by Gregory A. Petsko, Weill Cornell Medical College, New York, NY, and approved March 9, 2016 (received for review August 20, 2015)

It is well-established that widespread transcriptional changes accompany the onset and progression of Alzheimer's disease. Because of the multifactorial nature of this neurodegenerative disorder and its complex relationship with aging, however, it remains unclear whether such changes are the result of nonspecific dysregulation and multisystem failure or instead are part of a coordinated response to cellular dysfunction. To address this problem in a systematic manner, we performed a meta-analysis of about 1,600 microarrays from human central nervous system tissues to identify transcriptional changes upon aging and as a result of Alzheimer's disease. Our strategy to discover a transcriptional signature of Alzheimer's disease revealed a set of down-regulated genes that encode proteins metastable to aggregation. Using this approach, we identified a small number of biochemical pathways, notably oxidative phosphorylation, enriched in proteins vulnerable to aggregation in control brains and encoded by genes down-regulated in Alzheimer's disease. These results suggest that the down-regulation of a metastable subproteome may help mitigate aberrant protein aggregation when protein homeostasis becomes compromised in Alzheimer's disease.

neurodegenerative diseases | amyloid formation | protein misfolding | protein aggregation | protein supersaturation

Alzheimer's disease (AD) is a neurodegenerative condition responsible for the majority of reported cases of dementia, affecting over 44 million people worldwide (1–6). Although the exact nature of this disease has not been defined fully, its onset and progression have been associated with a multitude of factors, including mitochondrial dysfunction, disruption of the endoplasmic reticulum and membrane trafficking, disturbances in protein folding and clearance, and the activation of the inflammatory response (1–6). More generally, however, it is clear that AD belongs to a class of protein conformational disorders whose characteristic feature is that specific peptides and proteins misfold and aggregate to form amyloid assemblies (1, 3, 6). The presence of such aberrant aggregate species generates a cascade of pathological events, leading to the loss of the ability of protein homeostasis mechanisms to preserve normal biological function and to avoid the formation of toxic species (1, 3, 6).

The appearance of protein aggregates in living systems is increasingly recognized as being common, as growing evidence indicates that proteins are only marginally stable against aggregation in their native states (1, 7) and that the molecular processes that prevent protein aggregation decline with aging (8–12). Thus, protein aggregation is emerging as a widespread biological phenomenon, in which hundreds of different proteins can aggregate in aging, stress, or disease (9, 13–23). To understand why some proteins aggregate whereas others remain soluble, we recently observed that many proteins in the proteome are insufficiently soluble relative to their expression levels (24). Such proteins are metastable to aggregation as their concentrations exceed their solubilities, that is, they are supersaturated (24–27). Upon formation of aggregate seeds by nucleation events, a supersaturated protein will form insoluble deposits until the concentration of its soluble fraction is reduced to match its solubility (24–28). We found that the proteins that coaggregate with inclusion bodies, those that aggregate in aging, and those in the major biochemical pathways associated with neurodegenerative diseases tend to be supersaturated (24). The observation that these metastable proteins appear to be a common feature in aging, stress, and disease prompts the question of whether or not their supersaturation levels are altered in AD. These levels are particularly crucial, as supersaturation represents a major driving force for aggregation (25). It is thus interesting to ask whether the down-regulation of supersaturated proteins may limit their aggregation in response to compromised protein homeostasis.

In the present study, we examined the experimental information acquired in the last decade about transcriptional changes in AD (29–43). We aimed specifically to determine the relationship between protein supersaturation and the transcriptional changes that occur during normal aging and in AD. We found that distinct but partially overlapping transcriptional changes take place in aging and AD. Moreover, down-regulated genes generally correspond to metastable proteins at risk for aggregation, as they are supersaturated and encoded by highly expressed genes. Accordingly, the biochemical pathways down-regulated in AD are nearly identical to those previously identified as highly enriched in supersaturated proteins (24). These changes are also accompanied by a transcriptional down-regulation of certain components of the protein homeostasis network. The down-regulation

Significance

Alzheimer's disease, the most common cause of dementia, has been associated with a complex transcriptional response. To define the nature of this response, we carried out a comprehensive analysis that reveals a set of differentially expressed genes encoding proteins at risk for aggregation. These results identify a transcriptional signature of Alzheimer's disease consisting of down-regulated genes corresponding to a highly expressed "metastable subproteome" prone to aggregation. Our analysis of this metastable subproteome singles out a small number of biochemical pathways enriched in proteins that are simultaneously supersaturated in control brains and encoded by genes down-regulated in Alzheimer's disease.

Author contributions: P.C., R.K., R.I.M., C.M.D., and M.V. designed research; P.C., R.K., R.F., R.I.M., C.M.D., and M.V. performed research; P.C. and R.K. contributed new reagents/analytic tools; P.C., R.K., R.F., R.I.M., C.M.D., and M.V. analyzed data; and P.C., R.K., R.F., R.I.M., C.M.D., and M.V. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]P.C. and R.K. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: mv245@cam.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1516604113/-/DCSupplemental.

of genes corresponding to supersaturated proteins may thus represent a specific mechanism to limit widespread aggregation by regulating cellular concentrations in a compromised protein-folding environment.

Results

Analysis of the Transcriptional Changes in Aging and AD. A longstanding question is whether AD represents an acceleration of the normal aging process or a qualitatively distinct phenomenon. Determining changes in gene expression can offer important insights into this problem. The complications associated with obtaining human tissue samples, however, constrain the extent to which confounding variables such as age, gender, and tissue type can be controlled in a transcriptional analysis of AD. In the present work, the control samples (mean 70.8 ± 16.4 y) are younger than the disease samples (mean 81.1 ± 9.5 y), necessitating the use of techniques to account for these disparities (*SI Materials and Methods* and Table S1).

For the human genes examined in our analysis, we constructed a linear model of expression differences across a range of factors (*SI Materials and Methods*). We thus obtained the overall median magnitude and statistical significance of expression changes by combining these individual values across different studies. In this analysis, microarray probes were mapped onto UniProt IDs to determine the corresponding protein (*SI Materials and Methods*). Using this procedure, we determined the transcriptional changes associated with 19,254 genes. An important aspect of this approach is that the effects on gene expression of different factors are considered as additive. Because the occurrence of AD increases with age, Alzheimer's subjects exhibit specific disease-related transcriptional changes in addition to those associated with natural aging. We considered a gene to be differentially expressed if it undergoes a change in expression of at least 10% with a Benjamini–Hochberg–corrected P value ≤0.01. We then tested over 18,000 other combinations of thresholds and found our results to be robust to changes in these thresholds (Figs. S1 and S2). In the model used here the aging component is a linear variable, and therefore estimating the magnitude of change requires specifying a range of ages. Because the assumption of linearity is expected to hold best near the average age, we used the change in expression for an age range of approximately two SDs, namely 25 y.

Proteins That Aggregate in AD Correspond to Transcriptionally Down-Regulated Genes. We next asked how the transcriptional changes identified in aging and AD might be associated with protein aggregation. First, we considered the set of disease-related amyloid proteins, that is, those annotated as "amyloid" in UniProt, which include those associated with neurodegenerative diseases (24). On average, we could not detect an overall connection between amyloid proteins and proteins corresponding to differentially expressed genes (Fig. 1 *A* and *B*). We also note, however, that this analysis does not imply that individual genes in the amyloid class may not have important roles in AD. As an example, the down-regulation of the *APP* gene (in our analysis by 9.5%, with P =0.011) has been reported in neurons containing neurofibrillary tangles (44).

We identified, however, a clear signal for another set of proteins associated with AD, namely those that coaggregate with amyloid plaques (13) and neurofibrillary tangles (14) in human autopsy samples as identified by mass spectrometry. Among the proteins that coaggregate with plaques (35%, $P = 4.7{\cdot}10^{-3}$) and tangles (41%, $P = 1.7{\cdot}10^{-13}$), a disproportionate number correspond to down-regulated genes in AD (Fig. 1*A*) in addition to those that are down-regulated during natural aging (Fig. 1*C*). Proteins corresponding to genes down-regulated in aging are overrepresented among tangle coaggregators (10%, $P = 2.5{\cdot}10^{-3}$) but not plaque coaggregators (4%, $P = 1.0$) (Fig. 1*C*). By contrast, only an insignificant number of genes encoding proteins aggregating in plaques and tangles were observed to be up-regulated in AD (Fig. 1*B*) or aging (Fig. 1*D*).

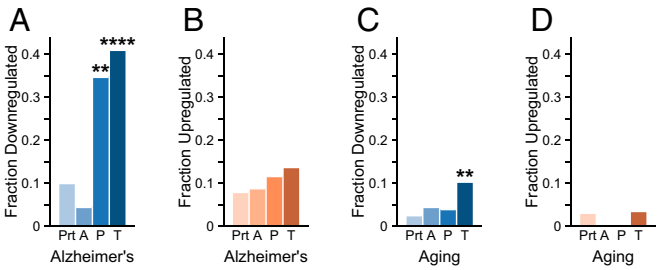

Fig. 1. Proteins that aggregate in AD correspond to transcriptionally down-regulated genes. (*A* and *B*) Fraction of proteins corresponding to transcriptionally down-regulated (*A*) or up-regulated (*B*) genes in AD in the whole proteome (Prt; down-regulated fraction 1,907/19,254; up-regulated fraction 1,509/19,254) and for amyloid deposits (A; 1/23; 2/23), plaques (P; 9/26; 3/26), and tangles (T; 36/88; 9/88). (*C* and *D*) Fraction of proteins corresponding to transcriptionally down-regulated (*C*) or up-regulated (*D*) genes in aging in the whole proteome (432/17,833; 534/17,833), and for amyloid deposits (1/23; 0/23), plaques (1/26; 0/26), and tangles (9/88; 3/88). The statistical significance of the difference with the proteome (first column) was assessed with a Fisher's exact test with Holm–Bonferroni corrections (**P < 0.001, ****P < 0.0001).

Metastable Proteins Correspond to Transcriptionally Down-Regulated Genes in Aging and AD. We next investigated whether the fact that so many proteins that coaggregate with plaques and tangles correspond to genes down-regulated in AD could be a consequence of their metastability to aggregation. We previously observed that these metastable proteins tend to be supersaturated, having concentrations exceeding their solubility limits (24). Here we calculated the metastability of proteins to aggregation in terms of supersaturation scores (σ_u), which represent the risk of proteins aggregating from their unfolded states (24). We assessed proteins corresponding to genes down-regulated in AD to be about 8.8-fold (8.8×, $P < 2.2{\cdot}10^{-16}$) more metastable than those for which the expression levels of the corresponding genes do not change significantly in disease (Fig. 2*A*). Similarly, we found proteins encoded by genes down-regulated in aging to be more metastable (7.4×, $P < 2.2{\cdot}10^{-16}$) than those whose expression does not change (Fig. 2*B*).

We also found that proteins corresponding to genes up-regulated in AD (1.3×, $P = 9.7{\cdot}10^{-13}$) (Fig. 2*A*) and in aging (1.5×, $P < 8.8{\cdot}10^{-7}$) (Fig. 2*B*) are modestly, but significantly, more metastable than those with unchanged expression in AD. These up-regulated genes are almost exclusively associated with an inflammatory response (Dataset S1). For example, of those genes that encode metastable proteins, the most highly up-regulated gene (123% increase in expression) in AD is alpha-1 antichymotrypsin, which inhibits serine proteases, particularly those active in inflammation (45).

Despite the fact that only 16% of down-regulated genes are common to aging and AD (Fig. 2*D*), in both cases the transcriptional response appears to be associated with metastability to aggregation (Fig. 2 *A*–*C*). Indeed, we observed a significant overlap ($P < 2.2{\cdot}10^{-16}$) between the most metastable proteins (≥95th percentile), proteins corresponding to genes down-regulated in AD, and proteins corresponding to genes down-regulated in aging, as well as between any two of these categories (Fig. 2*D*). The proteins that are supersaturated proteins and encoded by genes down-regulated in AD make up a metastable subproteome specific to AD (Dataset S1), which is here referred to as the "metastable subproteome." By contrast, the most transcriptionally up-regulated genes in AD and in aging overlap significantly with each other, but neither group is significantly enriched in genes encoding metastable proteins (Fig. 2*E*). As a control, we divided the down-regulated and up-regulated genes into low, medium, and high levels and calculated the supersaturation scores at each of these levels (Fig. 3). Our results indicated a trend toward increasing levels of supersaturation with increasing levels of down-regulation in AD (Fig. 3*A*). This correlation is weaker in

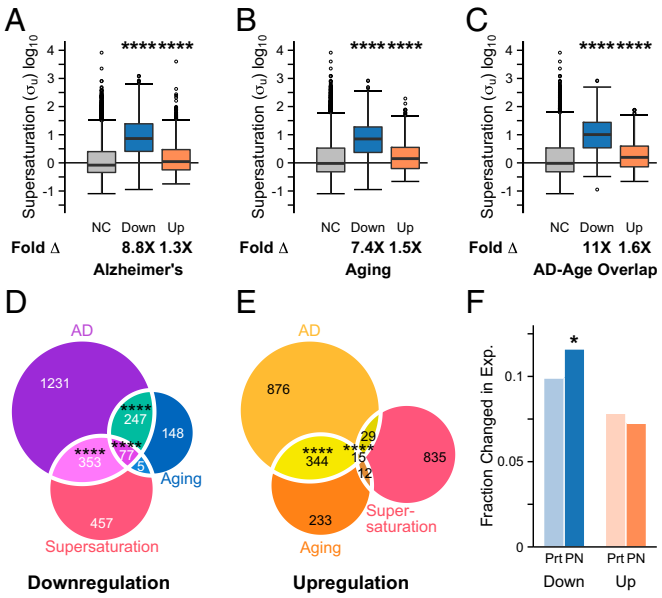

Fig. 2. Transcriptionally regulated genes in aging and AD correspond to proteins metastable against aggregation. (*A*–*C*) Assessment of the metastability to aggregation of the proteins associated with differentially expressed genes in (*A*) AD, (*B*) aging, and (*C*) the overlap between the two groups. The median fold difference in supersaturation (which is a measure of metastability to aggregation) is indicated by Fold Δ. NC, Down, and Up indicate, respectively, no change in expression, down-regulation, and up-regulation. Whiskers range from the lowest to highest value data points within 150% of the interquartile ranges. (*D* and *E*) Overlap between the 5% most supersaturated proteins and the corresponding genes either (*D*) down-regulated or (*E*) up-regulated in aging and AD. The number of proteins in each subset is indicated. (*F*) Fraction of genes down-regulated (blue) and up-regulated (orange) in the whole proteome (down-regulated fraction 1,907/19,254; up-regulated fraction 1,509/19,254) and the protein homeostasis network (PN; 1,509/19,254; 148/2,041). For *A*–*C*, $^{****}P \leq 0.0001$, one-sided Wilcoxon/Mann–Whitney test with Holm–Bonferroni correction. For *D* and *E*, $^{*}P \leq 0.05$, $^{****}P \leq 0.0001$, one-sided Fisher's exact test with Holm–Bonferroni correction.

aging (Fig. 3*C*), and weaker still among up-regulated genes (Fig. 3 *B* and *D*). The negative correlation between protein supersaturation and gene down-regulation also persists at the individual level for AD, but much less so for aging (Fig. S3).

Elevated supersaturation scores of differentially expressed genes may result from an easier detection of the differences in highly expressed genes than in genes of low expression. To control for this possibility, however, we excluded low-expression genes from our analysis, finding the median supersaturation of proteins corresponding to differentially expressed genes to be elevated even after this procedure (Fig. S4). We also tested the robustness of our results against changes in the details of our analysis. We found that our results on the metastability of the proteins corresponding to differentially expressed genes are stable across a wide range of thresholds for defining the groups of up-regulated and down-regulated genes (Figs. S1 and S2), and also against the introduction of Gaussian noise into the supersaturation score (Figs. S5 and S6).

Specific Protein Homeostasis Components Correspond to Genes Down-Regulated in AD. As we have discussed above, widespread down-regulation of genes corresponding to metastable proteins may represent a general mechanism to maintain protein homeostasis upon aging and AD. An additional transcriptional response, however, may also involve specific components of the protein homeostasis network (8). Following a recent study that showed an enrichment in genes down-regulated in aging in this network (8), we examined whether or not particular subnetworks in the overall protein homeostasis network correspond to genes particularly down-regulated in aging and AD (Fig. 2*F*). We found a significant number of protein homeostasis network genes in the "trafficking" subnetwork to be down-regulated in AD (14%, $P = 1.1 \cdot 10^{-2}$).

We then investigated whether or not the cell is endowed with transcriptional mechanisms to regulate the solubility burden in register with the protein homeostasis capacity. If so, there may be transcriptional regulators that coordinate such a response by modulating protein homeostasis. To determine in particular whether specific transcription factors and histone modifiers are up-regulated or down-regulated in AD and aging, we generated a map of transcriptional regulators and their targets using Encyclopedia of DNA Elements (ENCODE) regulator binding site data (46). Here we considered a gene to be regulated by a particular transcription factor or histone modifier if the regulator has a binding site at least half of which is within 1,000 bp of the start codon of the gene itself. We identified 23 transcription factors and histone modifiers associated with a significant number of genes down-regulated in AD (Dataset S2), including EGR1 (47), NRF1 (48), and REST (49). By contrast, we found only one regulator associated with a significant number of genes down-regulated in aging, the histone modifier EZH2 (Dataset S3). In addition, four regulators were found to be associated with a significant number of genes up-regulated in AD, and none was found to be associated with a significant number of genes up-regulated in aging (Datasets S2 and S3).

Biochemical Pathways Enriched in Metastable Proteins Are Also Enriched in Proteins Corresponding to Genes Down-Regulated in AD. To determine the functional implications of the transcriptional regulation of metastable proteins in AD, we conducted an unbiased search of the entire set of 284 pathways in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (50), a repository of biochemical pathways and protein networks. We found a close correspondence between the pathways down-regulated in AD

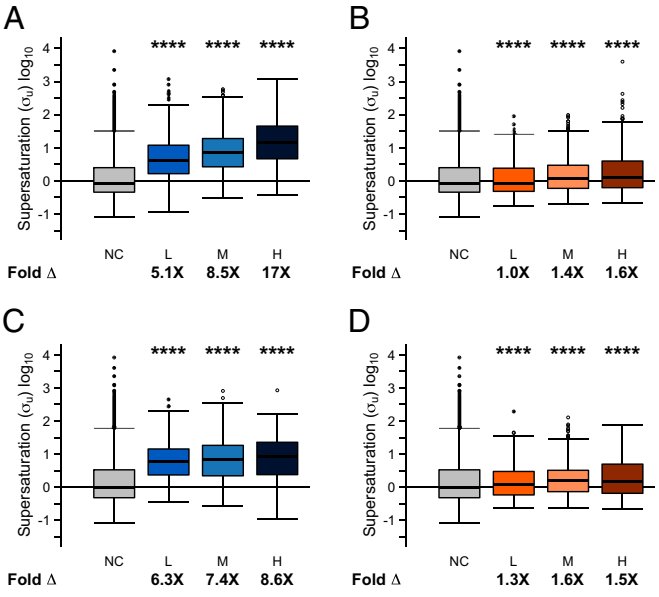

Fig. 3. Metastability of proteins to aggregation is correlated with the down-regulation of the corresponding genes in AD. Metastability levels, assessed by supersaturation scores, for proteins associated with differentially expressed genes: (*A*) down-regulated in AD, (*B*) up-regulated in AD, (*C*) down-regulated in aging, and (*D*) up-regulated in aging. Differentially expressed genes are divided into thirds (low, L; medium, M; high, H) based on the fold change of expression. The median fold difference in supersaturation is indicated by Fold Δ. NC indicates no change in expression. $^{****}P \leq 0.0001$, one-sided Wilcoxon/Mann–Whitney test with Holm–Bonferroni correction. Whiskers range from the lowest to highest value data points within 150% of the interquartile ranges.

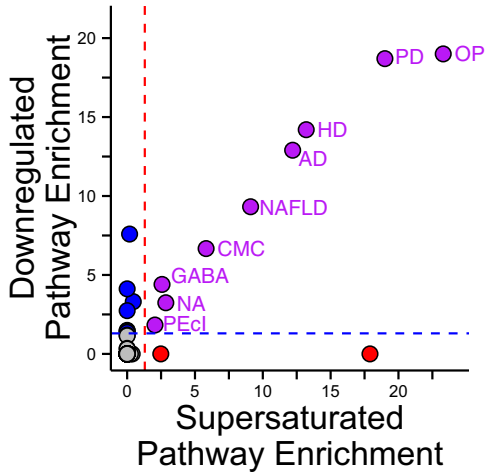

Fig. 4. Comparison between down-regulated and metastable biochemical pathways and networks. We found that the biochemical pathways and networks down-regulated in AD correspond closely to those enriched in supersaturated proteins (purple circles). Using the KEGG classification, these biochemical pathways and networks are oxidative phosphorylation, Parkinson's disease, Huntington's disease, Alzheimer's disease, nonalcoholic fatty liver disease, cardiac muscle contraction, nicotine addiction, GABAergic synapse, and pathogenic *E. coli* infection.

and pathways that we previously found to be supersaturated based on independent data (24, 25) (Fig. 4 and Table S2). Remarkably, most of these KEGG pathways fall along a band in which increasing metastability levels correspond to increasing down-regulation (Fig. 4, purple circles). The overlap between metastable and down-regulated pathways is highly significant ($P = 8.7{\cdot}10^{-11}$). Among the simultaneously metastable and down-regulated KEGG pathways, we found oxidative phosphorylation (OP), Parkinson's disease (PD), Huntington's disease (HD), Alzheimer's disease, nonalcoholic fatty liver disease (NAFLD), cardiac muscle contraction (CMC), nicotine addiction (NA), GABAergic synapse (GABA), and pathogenic *Escherichia coli* infection (PEcI). These results reveals pathological (AD, PD, HD, and NAFLD) and functional (OP, CMC, and PEcI) networks and pathways enriched in physiological complexes, as well as pathways involved in neuronal signaling (NA and GABA). In particular, our analysis identified certain proteins in the oxidative phosphorylation pathway as being particularly metastable, including all of the components of the mitochondrial ATP synthase complex for which we have data, consistent with the reported involvement of this complex in AD (51). In addition, 43% of the genes in our analysis that encode for mitochondrial ATP synthase complex are transcriptionally repressed. The most repressed is the alpha subunit of the F_1 catalytic core (whose expression is reduced by 26% in AD), which has been observed to accumulate in degenerating neurons in AD and to be associated with neurofibrillary tangles (52). We also verified that, although oxidative phosphorylation is central to the pathways down-regulated in AD, the signal for metastability in AD and aging is robust against the exclusion of proteins in this pathway from our analysis (Fig. S7).

In this comprehensive analysis of KEGG pathways, we also found other pathways that are significantly enriched in either metastable proteins (Fig. 4, red circles) or in proteins corresponding to down-regulated genes (Fig. 4, blue circles), but not both (Table S2). However, the large majority of these pathways have significance values that are lower than the average jointly metastable and down-regulated pathway (Fig. 4, purple circles), the exceptions being the "ribosome," which is highly metastable but not down-regulated, and the "synaptic vesicle cycle," "proteasome," and "retrograde endocannabinoid signaling," which are down-regulated but not metastable. A similar analysis for up-regulated pathways in AD did not provide particularly significant results, although one may expect genes associated with the immune response to be up-regulated, as, for example, complement C1q subcomponent subunit C and plasma protease C1 inhibitor in the "complement and coagulation cascade" pathway.

Thus, the observation that in AD there is a highly specific down-regulation of metastable biochemical pathways and networks suggests the presence of a robust transcriptional response to protein aggregation in AD.

Widespread Down-Regulation of the Metastable Subproteome Is Not a General Feature of Disease. Because the genes corresponding to the metastable subproteome are, on average, highly expressed, we considered the possibility that their widespread down-regulation could be a general feature of cellular dysfunction in disease. If this were the case, any process that disrupts normal cellular function could impair transcription, preferentially affecting those genes that are highly expressed. To investigate this possibility, we performed a meta-analysis of expression changes in another cognitive disorder, clinical depression. We considered 470 microarrays, including 239 from control patients and 231 from those with clinical depression (Table S1). As with our analysis of AD, we restricted our analysis to brain samples from cases in which the gender and age (for which we controlled) were known and Gene Expression Omnibus (GEO) database series that included at least 10 total cases. Among the 19,190 genes for which we evaluated changes in expression, we found 7 genes down-regulated and 11 genes up-regulated in clinical depression at the thresholds of 10% change in expression and $P \leq 0.01$ (Dataset S4). Overall, we did not observe the same widespread transcriptional repression of the metastable subproteome found in AD, and we found no KEGG pathways significantly enriched in proteins corresponding to those genes differentially expressed in AD.

We then considered the possibility that we had only identified a small number of genes as being differentially expressed in clinical depression because of low statistical power. Although our meta-analysis of clinical depression included only 22% as many arrays as that of AD, this is unlikely to explain the fact that only 0.6% as many genes are differentially regulated in clinical depression. In addition, our separate analysis for aging provided a control to assess the statistical power of the clinical depression dataset relative to that for AD. At the thresholds of 10% change in expression and $P \leq 0.01$, we found 196 genes down-regulated and 122 genes up-regulated in aging in the clinical depression dataset. This is 23% as many genes as we found differentially regulated in aging based on the AD dataset, consistent with the smaller number of microarrays in the clinical depression analysis. As a further control, we reanalyzed these data after relaxing the significance threshold for differential expression to $P \leq 0.05$. At this threshold, we found 24 genes down-regulated and 17 genes up-regulated in clinical depression and 569 genes down-regulated and 291 genes up-regulated in aging (Dataset S4). At the relaxed threshold, the KEGG pathway for "olfactory transduction" was enriched in proteins corresponding both to genes down-regulated ($P = 4.5{\cdot}10^{-3}$) and genes up-regulated ($P = 4.9{\cdot}10^{-2}$) in clinical depression (Dataset S4). Only "mineral absorption" was enriched in proteins corresponding to genes up-regulated in aging in the clinical depression dataset (Table S2). We also assessed the overall relationship between metastability and transcriptional regulation, and found little correlation between the two.

Discussion

A major area of investigation into the molecular origins of AD concerns the chemical and physical instability of the proteins associated with the disease, and the mechanisms by which the cell responds to such a situation. A number of studies have reported biophysical features, environmental conditions, and molecular partners that promote or repress the initial aggregation of specific proteins (1, 3, 7, 13–15). More recently, it has been recognized that the regulation of many other proteins is disrupted as a consequence of these initial aggregation events (8, 16–25). In a complementary approach, the origins of AD have been studied by analyzing

the transcriptional response associated with its onset and progression (29–43). These studies have revealed that this transcriptional response involves genes corresponding to proteins that can cause the disease and those associated with the cellular processes engaged in combating it (29–43).

In the present study, we have brought together these two approaches, finding that the transcriptional changes that occur in AD can be rationalized, at least in part, on the basis of the presence of an AD-specific metastable subproteome at risk for aggregation (Fig. 2). This metastable subproteome is defined as the overlap between the proteins that are most supersaturated and that correspond to highly expressed genes, and those encoded by genes most transcriptionally down-regulated in AD (Fig. 2*D*). These proteins are intrinsically at risk for aggregation and, as we found here, tend to be the target of the transcriptional response in aging and AD. These results are consistent with previous observations that the expression of oxidative phosphorylation genes is suppressed in AD (53, 54), but suggest in addition that such suppression may be part of a broader response to the disease.

Having previously shown that the proteins associated with AD tend to be metastable to aggregation because they are supersaturated (24, 25), we have now reported a response to this intrinsic metastability of the proteome in the face of disruptions to protein homeostasis through the transcriptional down-regulation of their respective genes. The close correspondence of the biochemical pathways associated with metastability and those down-regulated in AD (Fig. 4) supports this conclusion, as do the tendency for proteins that coaggregate in plaques and tangles to correspond to down-regulated genes (Fig. 1) and the high overall metastability level of proteins encoded by down-regulated genes (Fig. 2). We found these results to be stable against a range of potentially confounding factors, including the choice of thresholds for differential expression (Figs. S1 and S2), noise in the supersaturation score (Figs. S5 and S6), and the large contribution of oxidative phosphorylation (Fig. S7).

Analysis of the transcriptional response to the collapse of protein homeostasis in terms of a metastable subproteome at risk for aggregation has also enabled us to address another central question about the progression of AD, namely the way in which changes occurring in this disease are related to the natural process of aging. These results indicate that aging and AD are very different at the transcriptional level, as over three-quarters of the transcriptional changes that occur in AD do not occur in aging (Fig. 2 *D* and *E*). In addition, many cellular processes down-regulated in AD are not significantly down-regulated in aging (Fig. S2). Although the differences between regulation in aging and AD are profound, there are important commonalities, as shown by the significant overlap in the specific transcriptional changes that occur in AD and in aging (Fig. 3). AD therefore appears to involve an acceleration in the decline of protein homeostasis associated with aging, and also an extension of its scope and significance. Overall, such an acceleration makes the metastable subproteome that we have identified in this work more susceptible to aggregation. This conclusion offers an explanation of why a transcriptional down-regulation of genes corresponding to metastable proteins is observed in both aging and AD.

We also observe that these phenomena are unlikely to be a general feature of cellular dysfunction. Our results indicate that a different transcriptional response is present in the case of clinical depression (Table S2 and Dataset S4), consistent also with results for epilepsy derived considering the differentially expressed genes in hippocampal samples from five patients with mesial temporal lobe epilepsy with hippocampal sclerosis (55). In that study, 518 genes were found to be differentially expressed between the subjects. Functional enrichment using the Database for Annotation, Visualization and Integrated Discovery (DAVID, https://david.ncifcrf.gov/) showed enrichment for KEGG pathways associated with neuroactive ligand receptor interaction, drug metabolism, and cytokine interaction, among others. The KEGG pathways of oxidative phosphorylation and of Alzheimer's, Parkinson's, and Huntington's diseases were not, however, seen in epilepsy.

The findings that we have reported here, therefore, suggest that the widespread down-regulation of genes corresponding to metastable proteins at risk for aggregation may represent an important aspect of the strategy for cellular regulation in the face of disruptions in protein homeostasis. More generally, understanding the physicochemical implications of transcriptional regulation in aging, AD, and other protein-misfolding disorders has important implications both for a fundamental biological understanding of the origins of the disease and for clinical practice. Because the maintenance of protein homeostasis is an essential function in the cell, determining how the overall proteome composition is managed and modulated is a central question in biology. At the same time, understanding endogenous strategies for handling supersaturated, metastable, and potentially misfolding proteins may provide an avenue for improved therapies. If widespread aggregation is associated with AD, then determining how to regulate this phenomenon is of great value and practical importance.

Conclusions

We have shown that AD is associated with the transcriptional regulation of a metastable subproteome at risk for aggregation. The presence of these poorly soluble proteins in the cellular environment is inherently dangerous, in particular because these proteins tend to cluster into specific biochemical pathways, and only limited molecular chaperones and other protective resources are available at any given time to prevent their misfolding and aggregation. In conjunction with emerging insights into the molecular chaperone functions and the regulation of protein translation and degradation, our results indicate that the study of protein metastability may clarify how failures in maintaining proteins in their normal functional states could result in protein aggregation and in multifactorial disorders such as AD.

Despite the great complexity of aging processes and neurodegenerative disorders, protein solubility may underlie many aspects of their resultant cellular dysfunction. In this work, we have adopted this idea to investigate how the levels of poorly soluble proteins are regulated, finding that the overall transcriptional response to AD is associated with a global down-regulation of the expression of the genes encoding proteins that are metastable to aggregation. We anticipate that interventions that target the metastable subproteome at risk for aggregation that we have identified in this work may provide novel opportunities for the early diagnosis and treatment of AD.

Materials and Methods

The method of array normalization, construction of the linear model, and determination of significance and magnitude values are described in *SI Materials and Methods*. The calculation of basal expression levels for supersaturation scores and the sensitivity analysis are also described in *SI Materials and Methods*. The multiple hypothesis correction, KEGG analysis, and transcription factor analysis are described in *SI Materials and Methods*.

ACKNOWLEDGMENTS. P.C. was supported by grants from the US-UK Fulbright Commission, St. John's College, University of Cambridge, and the National Institutes of Health (Northwestern University Medical Scientist Training Program Grant T32 GM8152-28). R.I.M. was supported by grants from the National Institutes of Health (National Institute of General Medical Sciences, National Institute on Aging, National Institute of Neurological Disorders and Stroke), the Ellison Medical Foundation, the Glenn Foundation, and the Daniel F. and Ada L. Rice Foundation. C.M.D. and M.V. were supported by the Wellcome Trust.

1. Knowles TP, Vendruscolo M, Dobson CM (2014) The amyloid state and its association with protein misfolding diseases. *Nat Rev Mol Cell Biol* 15(6):384–396.
2. Querfurth HW, LaFerla FM (2010) Alzheimer's disease. *N Engl J Med* 362(4): 329–344.
3. Selkoe D, Mandelkow E, Holtzman D (2012) Deciphering Alzheimer disease. *Cold Spring Harb Perspect Med* 2(1):a011460.
4. Rubinsztein DC (2006) The roles of intracellular protein-degradation pathways in neurodegeneration. *Nature* 443(7113):780–786.
5. Glass CK, Saijo K, Winner B, Marchetto MC, Gage FH (2010) Mechanisms underlying inflammation in neurodegeneration. *Cell* 140(6):918–934.
6. Hardy J, Selkoe DJ (2002) The amyloid hypothesis of Alzheimer's disease: Progress and problems on the road to therapeutics. *Science* 297(5580):353–356.

7. Tartaglia GG, Pechmann S, Dobson CM, Vendruscolo M (2007) Life on the edge: A link between gene expression levels and aggregation rates of human proteins. *Trends Biochem Sci* 32(5):204–206.
8. Brehme M, et al. (2014) A chaperome subnetwork safeguards proteostasis in aging and neurodegenerative disease. *Cell Reports* 9(3):1135–1150.
9. Walther DM, et al. (2015) Widespread proteome remodeling and aggregation in aging *C. elegans*. *Cell* 161(4):919–932.
10. Ben-Zvi A, Miller EA, Morimoto RI (2009) Collapse of proteostasis represents an early molecular event in *Caenorhabditis elegans* aging. *Proc Natl Acad Sci USA* 106(35):14914–14919.
11. Labbadia J, Morimoto RI (2015) The biology of proteostasis in aging and disease. *Annu Rev Biochem* 84:435–464.
12. Labbadia J, Morimoto RI (2015) Repression of the heat shock response is a programmed event at the onset of reproduction. *Mol Cell* 59(4):639–650.
13. Liao L, et al. (2004) Proteomic characterization of postmortem amyloid plaques isolated by laser capture microdissection. *J Biol Chem* 279(35):37061–37068.
14. Wang Q, et al. (2005) Proteomic analysis of neurofibrillary tangles in Alzheimer disease identifies GAPDH as a detergent-insoluble paired helical filament tau binding protein. *FASEB J* 19(7):869–871.
15. Xia Q, et al. (2008) Proteomic identification of novel proteins associated with Lewy bodies. *Front Biosci* 13:3850–3856.
16. Olzscha H, et al. (2011) Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. *Cell* 144(1):67–78.
17. Gidalevitz T, Ben-Zvi A, Ho KH, Brignull HR, Morimoto RI (2006) Progressive disruption of cellular protein folding in models of polyglutamine diseases. *Science* 311(5766):1471–1474.
18. Chapman E, et al. (2006) Global aggregation of newly translated proteins in an *Escherichia coli* strain deficient of the chaperonin GroEL. *Proc Natl Acad Sci USA* 103(43):15800–15805.
19. David DC, et al. (2010) Widespread protein aggregation as an inherent part of aging in *C. elegans*. *PLoS Biol* 8(8):e1000450.
20. Reis-Rodrigues P, et al. (2012) Proteomic analysis of age-dependent changes in protein solubility identifies genes that modulate lifespan. *Aging Cell* 11(1):120–127.
21. Koga H, Kaushik S, Cuervo AM (2011) Protein homeostasis and aging: The importance of exquisite quality control. *Ageing Res Rev* 10(2):205–215.
22. Koplin A, et al. (2010) A dual function for chaperones SSB-RAC and the NAC nascent polypeptide-associated complex on ribosomes. *J Cell Biol* 189(1):57–68.
23. Narayanaswamy R, et al. (2009) Widespread reorganization of metabolic enzymes into reversible assemblies upon nutrient starvation. *Proc Natl Acad Sci USA* 106(25):10147–10152.
24. Ciryam P, Tartaglia GG, Morimoto RI, Dobson CM, Vendruscolo M (2013) Widespread aggregation and neurodegenerative diseases are associated with supersaturated proteins. *Cell Reports* 5(3):781–790.
25. Ciryam P, Kundra R, Morimoto RI, Dobson CM, Vendruscolo M (2015) Supersaturation is a major driving force for protein aggregation in neurodegenerative diseases. *Trends Pharmacol Sci* 36(2):72–77.
26. Hofrichter J, Ross PD, Eaton WA (1976) Supersaturation in sickle cell hemoglobin solutions. *Proc Natl Acad Sci USA* 73(9):3035–3039.
27. Ikenoue T, et al. (2014) Heat of supersaturation-limited amyloid burst directly monitored by isothermal titration calorimetry. *Proc Natl Acad Sci USA* 111(18):6654–6659.
28. Muta H, et al. (2014) Supersaturation-limited amyloid fibrillation of insulin revealed by ultrasonication. *J Biol Chem* 289(26):18228–18238.
29. Ho L, et al. (2001) Altered expression of a-type but not b-type synapsin isoform in the brain of patients at high risk for Alzheimer's disease assessed by DNA microarray technique. *Neurosci Lett* 298(3):191–194.
30. Blalock EM, et al. (2004) Incipient Alzheimer's disease: Microarray correlation analyses reveal major transcriptional and tumor suppressor responses. *Proc Natl Acad Sci USA* 101(7):2173–2178.
31. Umemura K, et al. (2006) Autotaxin expression is enhanced in frontal cortex of Alzheimer-type dementia patients. *Neurosci Lett* 400(1-2):97–100.
32. Liang WS, et al. (2007) Gene expression profiles in anatomically and functionally distinct regions of the normal aged human brain. *Physiol Genomics* 28(3):311–322.
33. Liang WS, et al. (2008) Alzheimer's disease is associated with reduced expression of energy metabolism genes in posterior cingulate neurons. *Proc Natl Acad Sci USA* 105(11):4441–4446.
34. Webster JA, et al.; NACC-Neuropathology Group (2009) Genetic control of human brain transcript expression in Alzheimer disease. *Am J Hum Genet* 84(4):445–458.
35. Tan MG, et al. (2010) Genome wide profiling of altered gene expression in the neocortex of Alzheimer's disease. *J Neurosci Res* 88(6):1157–1169.
36. Simpson JE, et al.; MRC Cognitive Function and Ageing Neuropathology Study Group (2011) Microarray analysis of the astrocyte transcriptome in the aging brain: Relationship to Alzheimer's pathology and APOE genotype. *Neurobiol Aging* 32(10):1795–1807.
37. Cooper-Knock J, et al. (2012) Gene expression profiling in human neurodegenerative disease. *Nat Rev Neurol* 8(9):518–530.
38. Durrenberger PF, et al. (2012) Selection of novel reference genes for use in the human central nervous system: A BrainNet Europe study. *Acta Neuropathol* 124(6):893–903.
39. Antonell A, et al. (2013) A preliminary study of the whole-genome expression profile of sporadic and monogenic early-onset Alzheimer's disease. *Neurobiol Aging* 34(7):1772–1778.
40. Hokama M, et al. (2014) Altered expression of diabetes-related genes in Alzheimer's disease brains: The Hisayama study. *Cereb Cortex* 24(9):2476–2488.
41. Miller JA, Woltjer RL, Goodenbour JM, Horvath S, Geschwind DH (2013) Genes and pathways underlying regional and cell type changes in Alzheimer's disease. *Genome Med* 5(5):48.
42. Zhang B, et al. (2013) Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell* 153(3):707–720.
43. Ding B, et al. (2014) Gene expression profiles of entorhinal cortex in Alzheimer's disease. *Am J Alzheimers Dis Other Demen* 29(6):526–532.
44. Ginsberg SD, Hemby SE, Lee VM, Eberwine JH, Trojanowski JQ (2000) Expression profile of transcripts in Alzheimer's disease tangle-bearing CA1 neurons. *Ann Neurol* 48(1):77–87.
45. Baker C, Belbin O, Kalsheker N, Morgan K (2007) SERPINA3 (aka alpha-1-antichymotrypsin). *Front Biosci* 12:2821–2835.
46. ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489(7414):57–74.
47. Koldamova R, et al. (2014) Genome-wide approaches reveal EGR1-controlled regulatory networks associated with neurodegeneration. *Neurobiol Dis* 63:107–114.
48. Sheng B, et al. (2012) Impaired mitochondrial biogenesis contributes to mitochondrial dysfunction in Alzheimer's disease. *J Neurochem* 120(3):419–429.
49. Lu T, et al. (2014) REST and stress resistance in ageing and Alzheimer's disease. *Nature* 507(7493):448–454.
50. Whitlock MC (2005) Combining probability from independent tests: The weighted Z-method is superior to Fisher's approach. *J Evol Biol* 18(5):1368–1373.
51. Terni B, Boada J, Portero-Otin M, Pamplona R, Ferrer I (2010) Mitochondrial ATP-synthase in the entorhinal cortex is a target of oxidative stress at stages I/II of Alzheimer's disease pathology. *Brain Pathol* 20(1):222–233.
52. Sergeant N, et al. (2003) Association of ATP synthase α-chain with neurofibrillary degeneration in Alzheimer's disease. *Neuroscience* 117(2):293–303.
53. Chandrasekaran K, Hatanpää K, Rapoport SI, Brady DR (1997) Decreased expression of nuclear and mitochondrial DNA-encoded genes of oxidative phosphorylation in association neocortex in Alzheimer disease. *Brain Res Mol Brain Res* 44(1):99–104.
54. Manczak M, Park BS, Jung Y, Reddy PH (2004) Differential expression of oxidative phosphorylation genes in patients with Alzheimer's disease: Implications for early mitochondrial dysfunction and oxidative damage. *Neuromolecular Med* 5(2):147–162.
55. Johnson MR, et al. (2015) Systems genetics identifies Sestrin 3 as a regulator of a proconvulsant gene network in human epileptic hippocampus. *Nat Commun* 6:6031.
56. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol* 57(1):289–300.
57. Fisher RA (1925) *Statistical Methods for Research Workers* (Oliver and Boyd, London).
58. Pearson ES (1938) The probability integral transformation for testing goodness of fit and combining independent tests of significance. *Biometrika* 30(1-2):134–148.
59. Stouffer SA, Suchman EA, DeVinney LC, Star SA, Williams RM, Jr (1949) *The American Soldier: Adjustment During Army Life* (Princeton Univ Press, Princeton, NJ).
60. Lanz TA, et al. (2015) STEP levels are unchanged in pre-frontal cortex and associative striatum in post-mortem human brain samples from subjects with schizophrenia, bipolar disorder and major depressive disorder. *PLoS One* 10(3):e0121744.
61. Holm S (1979) A simple sequentially rejective multiple test procedure. *Scand J Stat* 6(2):65–70.
62. Iwamoto K, Kakiuchi C, Bundo M, Ikeda K, Kato T (2004) Molecular characterization of bipolar disorder by comparing gene expression profiles of postmortem brains of major mental disorders. *Mol Psychiatry* 9(4):406–416.
63. Chang L-C, et al. (2014) A conserved BDNF, glutamate- and GABA-enriched gene module related to human depression identified by coexpression meta-analysis and DNA variant genome-wide association studies. *PLoS One* 9(3):e90980.
64. Duric V, et al. (2010) A negative regulator of MAP kinase causes depressive behavior. *Nat Med* 16(11):1328–1332.