A Fragment-Based Method of Creating Small-Molecule Libraries to Target the Aggregation of Intrinsically Disordered Proteins

Priyanka Joshi, Sean Chia, Johnny Habchi, Tuomas P. J. Knowles, Christopher M. Dobson, and Michele Vendruscolo*  
Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom  
Supporting Information

ABSTRACT: The aggregation process of intrinsically disordered proteins (IDPs) has been associated with a wide range of neurodegenerative disorders, including Alzheimer’s and Parkinson’s diseases. Currently, however, no drug in clinical use targets IDP aggregation. To facilitate drug discovery programs in this important and challenging area, we describe a fragment-based approach of generating small-molecule libraries that target specific IDPs. The method is based on the use of molecular fragments extracted from compounds reported in the literature to inhibit the aggregation of IDPs. These fragments are used to screen existing large generic libraries of small molecules to form smaller libraries specific for given IDPs. We illustrate this approach by describing three distinct small-molecule libraries to target, A\textsubscript{β}, tau, and \(\alpha\)-synuclein, which are three IDPs implicated in Alzheimer’s and Parkinson’s diseases. The strategy described here offers novel opportunities for the identification of effective molecular scaffolds for drug discovery for neurodegenerative disorders and to provide insights into the mechanism of small-molecule binding to IDPs.

KEYWORDS: protein aggregation, drug discovery, Alzheimer’s disease, Parkinson’s disease

INTRODUCTION

Neurodegenerative disorders, including Alzheimer’s and Parkinson’s diseases, are highly debilitating conditions that progressively impair cognitive and motor skills with eventually fatal consequences.\textsuperscript{1–4} Such disorders already have the largest healthcare and social costs in the modern world, and their impact is set to rise even further in the future with the aging of the population.\textsuperscript{1,3} Despite over 20 years of intense research, however, no disease-modifying drug has yet entered clinical use.\textsuperscript{5–7}

This situation can be attributed, at least in part, to an incomplete understanding of the fundamental origins of such diseases. A range of factors have been implicated in their onset and progression, including the deficiencies in the protein production and degradation mechanisms, mitochondrial dysfunction, oxidative stress, the disruption of the endoplasmic reticulum and of membrane trafficking, and the activation of inflammatory responses.\textsuperscript{1,3,6–10} Despite this tremendous complexity, a common feature shared among these disorders is that specific peptides and proteins, including A\textsubscript{β} and tau in Alzheimer’s disease and \(\alpha\)-synuclein in Parkinson’s disease, aggregate into amyloid assemblies.\textsuperscript{1,2,4,11} The presence of such aberrant aggregates can trigger a cascade of pathological events, leading to the progressive failure of protein homeostasis and the irreversible loss of normal biological function.\textsuperscript{1,2,4,12–14}

Targeting the aggregation process of the peptides and proteins associated with neurodegenerative disorders has therefore been a primary route for drug discovery, and a range of inhibitors of A\textsubscript{β}, tau, and \(\alpha\)-synuclein aggregation have been reported in the literature.\textsuperscript{15–28} In the absence of successful compounds at the clinical stage, however, there is a pressing need to develop new strategies for the identification of more effective inhibitors of the aggregation process of these proteins. A major complication to achieve this goal arises from the fact that, in contrast to ordered proteins, whose structures exhibit only small conformational fluctuations, intrinsically disordered
proteins (IDPs) exist as ensembles of structures with highly heterogeneous conformational properties. In particular, in the absence of accurate structures or ensembles of structures, it is highly challenging to follow rational structure-based approaches, where small molecules are designed to be complementary in shape and charge to the target with which they are meant to interact and subsequently bind. An alternative approach is to screen large sets of compounds from existing small-molecule libraries against the IDPs involved in neurodegenerative disorders. Such an approach, however, is time-consuming and costly, not least because many IDPs are aggregation prone and difficult to purify. Moreover, screening large libraries does not give readily rational insights into the chemical space explored by the potential inhibitors of a given target.

A further complication is that there are currently no target-specific libraries available for IDPs. This situation is not ideal; as in the case of other major diseases, great advances have been made by using open access and commercial drug databases consisting of specific compounds for particular drug targets, including for example G protein-coupled receptors and protein kinases. In drug screening against a given target, chemicals in these databases cover a substantial portion of the chemical space with features specific for that target. To cover this gap and build IDP-specific small-molecule libraries, we report here a fragment-based strategy (Figure 1) that exploits the knowledge of reported inhibitors of IDP aggregation to identify a wide range of new potential inhibitors from open access compound repositories. Our overall goal is to use the chemical knowledge from reported inhibitors to identify from large databases additional potential inhibitors of IDP aggregation. A further objective is then to map a range of different pharmacophoric elements required for high binding. This overall strategy, which we illustrate in this work by creating three distinct libraries to target, Aβ, tau, and α-synuclein, is general and can be applied to design small-molecule libraries for other IDPs.

**RESULTS**

**A Small-Molecule Library against Aβ Aggregation.**

Compilation of an Initial Set of Compounds Reported to

Inhibit Aβ Aggregation. Several compounds have been reported to bind different structural species of Aβ, including monomers, oligomers, and amyloid fibrils. We mined this information in the literature and assembled an initial set of 88 compounds (see Materials and Methods), although in several cases no details were available regarding the specific species to which they bind. Two further compounds (retinoic acid and bromocriptine) with IC50 values in the range 10–20 μM were also considered to incorporate fragment diversity. These compounds were divided into categories with respect to the structural species of Aβ with which they interact (Figure 2).

Only five compounds were reported to bind Aβ monomers. Because binding the monomeric form of IDPs and stabilizing it is an attractive therapeutic strategy, these chemical scaffolds can provide insights into the chemical moieties that may interact with Aβ. To maintain generality in targeting the Aβ aggregation process, however, as mentioned above, compounds that interact with other species, namely oligomers and fibrils, were also considered.
Generation of a Set of Molecular Fragments Specific to Modulate Aβ Aggregation. A set of 164 fragments was obtained from the 88 compounds in the initial set described above. These fragments are expected to have chemical features capable of modulating the aggregation behavior of Aβ, as they are derived from reported inhibitors (see Materials and Methods). We screened these fragments with public ligand repositories, namely ZINC, ChEMBL, PubChem, and DrugBank (Figure 3). These databases contain curated small molecules with corresponding physicochemical properties and bioactivity data. A total of 16,850 compounds were obtained (Figure 4). These compounds comprise drugs in various categories, including antibiotics, nutraceuticals, and natural compounds. Compounds in untested and dubious categories were filtered out.

Analysis of the Compounds in the Aβ Small-Molecule Library. After having generated a small-molecule library to target Aβ aggregation, we analyzed the compounds present in it to identify possible characteristic chemical features. Because most of the inhibitors of Aβ in the initial set of 88 compounds from the literature are derived from polyphenols, the catechol substructure was the most common fragment observed in the databases (Figure 5), suggesting that the phenyl group with the hydroxyl moiety may have significant interactions with Aβ monomers, or with any of its assemblies, and play a role in modulating its aggregation behavior.

It has also been reported in structure–activity relationship (SAR) studies of curcumin that an aromatic ring with terminal hydroxyl substitutions is important for the activity of an effective inhibitor of aggregation. This particular substructure...
is present for instance in resveratrol, a reported inhibitor where a terminal phenyl group with two hydroxyl substitutions is present. Apart from potential antioxidant activities, resveratrol may directly bind to Aβ42, interfere in Aβ42 aggregation, change the Aβ42 oligomer conformation, and attenuate Aβ42 oligomeric cytotoxicity.24,49

Taken together, these observations indicate that compounds that target Aβ aggregation tend to show a terminal phenyl group and a hydroxyl substitution on the aromatic end group (Figure 6), as in the compounds reported in the literature, it was seen that loss of the hydroxyl group on the aromatic rings abolished inhibitory activity. In addition, the length of the linker tends to be between 8 and 16 Å (Figure 6). This
observation is supported by a hotspot analysis of small-molecule binding pockets in the soluble monomeric form of Aβ42 peptide where the distance between any of the two binding pockets is in the given range. A compound in the library may have any of the two fragments from known inhibitors connected by a linker of appropriate length. Additionally, the linker tends to be rigid with less than one to two freely rotating bonds, although some reported compounds violate this condition. It will be interesting to test in vitro which of these observations and in what combinations are suitable for an effective inhibitor. These three criteria are consistent with previous studies where similar features were observed for the inhibitors.19

A Small-Molecule Library against Tau Aggregation. Compilation of an Initial Set of Compounds Reported to Inhibit Tau Aggregation. As described above in the case of Aβ, for tau the first step in our approach also consisted of deriving from the existing literature an initial set of compounds active against tau aggregation (see Materials and Methods). Quite generally, current anti-tau therapeutic strategies have been aimed at identifying three major types of molecules: (1) inhibitors of protein kinases and phosphatases that modify the hyperphosphorylation of tau associated with its cellular aggregation, (2) compounds related to methylene blue, and (3) natural phytocomplexes and polyphenolic compounds able to inhibit the formation of tau filaments or to disaggregate them. We focus here on the third type of molecules with the overall aim to identify molecules that form favorable interactions with tau. We thus obtained a set of 23 small molecules reported to inhibit tau aggregation (Figure 4), including phenothiazines,50 cyanine dyes,51 polyphenol,52 phthalocyanine,53 anthraquinone,54 N-phenylamines,54 phenyl-thiazolyl-hydrazides,55 rhodanines,56,57 quinoxalines,58 and thienopyridazines.59 These compounds represent a wide variety

Figure 7. Compounds selected from the Aβ library are able to inhibit Aβ42 aggregation in vitro in a concentration-dependent manner. Ferriprotoporphyrin IX decreased the amount of fibrils by approximately 98% at one molar equivalent of Aβ42 compared to that of eriochrome blue black B, which decreased the amount of fibrils by approximately 73% at the same concentration.
of chemical structures but share the common feature of having multiple aromatic rings.

**Generation of a Set of Fragments Specific to Modulate Tau Aggregation.** We assumed that the 23 compounds described above have chemical substructures capable of modulating the aggregation of tau. In the second step of our procedure, we thus used these compounds to derive 57 fragments with tau binding properties (see Materials and Methods and Figure 4).

**Screening of Fragments to Identify Tau Aggregation Inhibitors.** In the third step of our procedure, we screened the fragments described above with existing large repositories of small molecules (see Materials and Methods). A total of 11,800 compounds were thus obtained, which constitute our library of predicted tau aggregation inhibitors (Figure 4).

**A Small-Molecule Library against α-Synuclein Aggregation. Compilation of an Initial Set of Compounds Reported to Inhibit α-Synuclein Aggregation.** Following the approach described above for Aβ and tau, in the first step of our procedure (see Materials and Methods), we obtained from the literature a set of 72 small molecules reported to bind α-synuclein as seed compounds (Figure 4).

**Generation of a Set of Fragments Specific to Modulate α-Synuclein Aggregation.** From the 72 compounds in the list presented above, we obtained 139 fragments (Figure 4) that are expected to possess chemical substructures capable of modulating α-synuclein aggregation behavior (see Materials and Methods).

**Screening of Fragments to Identify α-Synuclein Aggregation Inhibitors.** In the third step of our procedure, we then screened these 139 fragments with the existing large-molecule repositories (see Materials and Methods). These databases contain curated small molecules with their physicochemical properties and bioactivity data. A total of 14,735 compounds were thus obtained (Figure 4).

**Experimental Validation of the Compounds in the Libraries.** To investigate the richness of the small-molecule library in terms of active compounds, we selected 15 compounds with different chemical scaffolds from the Aβ library and tested them experimentally for their effect on Aβ aggregation. This procedure not only helps us gain insight into the quality of the library but also into the type of chemical scaffolds that can lead to the generation of more effective libraries in the future. Out of these 15 compounds, three decreased Aβ aggregation (bexarotene, eriochrome blue black E, and protoporphyrin IX), seven did not affect Aβ aggregation (bosentinib, glutamine, 4-aminosalicylic acid, colchicine, EGCG, scylo-inositol, and PADK), one increased Aβ aggregation (myricetin), and four showed weak modulatory effects on Aβ aggregation (Z-pro-prolinol, melofenamic acid, isothearmine, and linezolid). From these results, one could estimate a 20% success rate.

To further illustrate these results, we describe the results for the three compounds that we found to be active against Aβ aggregation. These compounds were tested using a highly reproducible thioflavin-T (ThT)-based protocol as described previously. The ThT-based assay allows assessing the effect of small molecules on the aggregation kinetics of the Aβ42 peptide by binding to any of the coexisting pathway Aβ42 species. The first small molecule is eriochrome blue black B, which is a structural derivative of the complexometric indicator eriochrome black T. The second compound is ferrprotoporphyrin IX, which is a natural compound released when heme is liberated from hemoglobin in the presence of oxygen. These compounds are the result of applying the fragment-based approach described in the current strategy on Congo red and protoporphyrin IX. Indeed, they possess all the chemical features that we identified as characteristic for small molecules to bind Aβ42. Subsequently, Aβ42 fibril formation was monitored in the absence and presence of both compounds for 15 h. We found that these two compounds are able to inhibit Aβ42 aggregation substantially in a concentration-dependent manner (Figure 7). The values of ThT fluorescence that reflect fibril mass concentration were substantially decreased in the presence of the compounds compared to that of Aβ42 alone with this decrease being even more pronounced for ferrprotoporphyrin IX. Indeed, this latter molecule decreased the amount of fibrils by approximately 98% at one molar equivalent of Aβ42 compared to that of eriochrome blue black B, which decreased the amount of fibrils by approximately 73% at the same concentration (Figure 7). Irrespective of the extent of the effect of these molecules in Aβ42 aggregation, these results indicate that the proposed fragment-based approach has led to the identification of small molecules that are able to inhibit Aβ42 aggregation. The third compound, bexarotene, which is a retinoid X receptor (RXR) agonist approved by the FDA for the treatment of cutaneous T cell lymphoma, was also shown to be able to suppress the formation of toxic species in neuroblastoma cells and in a C. elegans model of Alzheimer’s disease.

Taken together, these results indicate that the libraries that we describe in this work may represent a useful resource to identify novel compounds against protein aggregation.

**DISCUSSION**

**Toward a Rational Design of Small-Molecule Inhibitors of IDP Aggregation.** As candidate drugs for neurodegenerative disorders are failing in clinical trials, there is a pressing need to not only identify new successful lead molecules but also elucidate the reasons for the failures.

The three small-molecule libraries described in this work to target Aβ, tau, and α-synuclein aggregation could offer novel opportunities for further investigation into these issues. We have provided an initial analysis in this direction by describing three characteristic features of Aβ inhibitors (Figure 6). These types of studies could eventually lead to general rational design rules for targeting IDPs and help increase our understanding of the chemical space of small molecules capable of interacting with IDPs. We anticipate that in further studies it will be possible to carry out structure–activity relationship (SAR) analyses of the compounds described here using various biophysical and chemical kinetics approaches to further clarify their chemical features and mechanisms of action.

**Conformational Species Targeted by the Small Molecules Described in This Work.** When developing inhibitors of IDP aggregation, a major question concerns the types of conformations that can be targeted, including monomers, oligomers, fibrillar intermediates, and mature amyloid fibrils. Although soluble oligomers are known to contribute mostly to toxicity, binding to other species may also interfere with the aggregation process and lead to successful interventions. Quite generally, we anticipate that the fragments and compounds described in this work could be further modified and linked in various permutations to obtain higher affinity ligands having the potential to bind the pathological species of interest.
Drug Repurposing. The three fragment-based libraries described in this work consist of small molecules with up to three fragments extracted from compounds reported in the literature to bind the target IDPs. From these libraries, one can identify small molecules to be developed into lead molecules for Alzheimer’s and Parkinson’s diseases. As some of the compounds in these libraries have other known targets, they may be considered as candidates for drug repurposing. Because a considerable amount of time and resources has already been devoted to their development for other diseases,65,66 bringing them to clinical trials for Alzheimer’s and Parkinson’s diseases could reduce the burden on research and development (Figure 8).

Figure 8. FDA-approved drugs in the three small-molecule libraries described in this work. The DrugBank contains 7740 drug entries, including 1584 FDA-approved small-molecule drugs. From this list, we identified 386 FDA-approved drugs predicted to inhibit Aβ aggregation, 392 for α-synuclein, and 467 for tau.

CONCLUSIONS

We have presented an approach to generate small-molecule libraries with specific antiaggregation activity against target IDPs. We have illustrated this approach by reporting three small-molecule libraries against Aβ, tau, and α-synuclein, which are three IDPs implicated in Alzheimer’s and Parkinson’s diseases. These small molecules, together with an analysis of their properties, could offer novel starting points for the development of therapeutic leads against neurodegenerative diseases and reveal general principles for the rational design of small molecules to target IDP aggregation.

MATERIALS AND METHODS

Creation of Libraries of Small Molecules to Target Given IDPs. The fragment-based drug discovery procedure that we describe in this work is divided into three steps.

Step 1: Assembly of an Initial Set of Compounds from Literature Mining. The first step in our approach is to create an initial set of compounds already reported in the literature to inhibit the aggregation of a target IDP (Figure 1). In this work, because we are concerned with three different IDPs, Aβ, α-synuclein, and tau, we built three distinct initial sets of compounds. We describe here the procedure in the case of Aβ, as similar steps were used for α-synuclein and tau. To collect compounds from the literature, we performed literature text mining in PubMed and Google Scholar to retrieve research articles involving the following keywords: “Aβ”, “Alzheimer’s disease”, “aggregation”, “monomer”, “oligomer”, and “fibris”.

We further manually sorted out this initial set to find mechanisms of action and IC₅₀ values, if described, of the compounds reported. Manual curation was essential to select only those compounds reported to bind to structural species associated with the aggregation process. In particular, we did not include acetylcholinesterase inhibitors, NMDA receptor antagonists, peptides, and monoclonal antibodies even though they have effects on aggregation through other mechanisms. Compounds showing unclear binding data were also discarded. The resulting compounds were further selected according to the two following conditions: (i) IC₅₀ values in the range 0.012–3.2 μM and (ii) inhibition of the aggregation of Aβ. In many cases, compounds were reported to also bind α-synuclein and tau; we selected compounds of this type reported to have an IC₅₀ value ≤ 10 μM.

The step 1 compounds for Aβ, α-synuclein, and tau are provided in Table S1.

Step 2: Generation of a Set of Fragments from the Initial Set of Compounds. In the second step, we use the known compounds in the initial set generated in step 1 as the starting points or “seeds”. For these seeds, in the SMILES format, we calculated pairwise Tanimoto similarity scores67 to avoid the presence of compound pairs exceeding 0.75 in similarity score in the set to be carried forward. Each compound was then cleaved into fragments by cutting acyclic bonds, and the fragments were added to the set of fragments. The features of the resulting fragments that were retained are (i) molecular weight ≤ 300 Da, (ii) hydrogen bond acceptors ≤ 3, and (iii) hydrogen bond donors ≤ 3. These features are selected because they are expected to give rise to druglike compounds.

Step 3: Generation of the Small-Molecule Library by a Screening of the Fragments. The set of fragments generated in step 2 is used to screen large databases of small molecules. The fragments were imported as “data structure” into ChemAxon (http://www.chemaxon.com). Instant JChem (http://www.chemaxon.com) was used for structure database management, search, and prediction. Each fragment was then iteratively screened with four large small-molecule databases, ZINC,43 ChEMBL,44 PubChem,45 and DrugBank.46 This procedure gave us compounds from each database. Redundancy was resolved by removing overlapping compounds. This list of compounds was further reduced in number by applying Lipinski’s ”rule of 5”68.

The libraries of compounds are given in Table S2 (Aβ), Table S3 (α-synuclein), and Table S4 (tau).

Step 4: In Vitro Validation of Compounds in the Aβ Library. Preparation of Aβ Peptides. The 42-residue form of Aβ peptide, here called Aβ42, with sequence MDAEFRHDS-GYEVHHQKLFFAEVGSGKIIIGLMVGGYVIA, was expressed in the E. coli BL21 Gold (DE3) strain (Stratagene, CA, U.S.A.) and purified as described previously with slight modifications.45 Briefly, the purification procedure involved sonication of E. coli cells, dissolution of inclusion bodies in 8 M urea, and ion exchange in batch mode on diethylaminoethyl cellulose resin and lyophilization. The lyophilized fractions were further purified using Superdex 75 HR 26/60 column (GE Healthcare, Buckinghamshire, UK), and eluates were analyzed using SDS-PAGE for the presence of the desired protein product. The fractions containing the recombiant protein were combined, frozen using liquid nitrogen, and lyophilized again. Chemicals, including eriochrome blue black B and
ferriprotoporphyrin IX, were obtained from Sigma-Aldrich and were of the highest purity available.

Preparation of Peptide Samples for Thioflavin T (ThT) Experiments. Solutions of monomeric Aβ42 peptides were prepared by dissolving lyophilized peptides in 6 M GuHCl. Monomeric forms were purified from potential oligomeric species and salt using a Superdex 75 10/300 GL column (GE Healthcare) at a flow rate of 0.5 mL/min and were eluted in 20 mM sodium phosphate buffer, pH 8, supplemented with 200 μM EDTA and 0.02% NaN3. The center of the peak was collected, and the peptide concentration was determined from the absorbance of the integrated peak area using ε280 = 1400 L mol⁻¹ cm⁻¹. The obtained monomer was diluted with buffer to the desired concentration and supplemented with 20 μM ThT from a 2 mM stock. All samples were prepared in low binding Eppendorf tubes on ice using careful pipetting to avoid introduction of air bubbles. Each sample was then pipetted into multiple wells of a 96-well half-area, low-binding, clear-bottom, PEG-coated plate (Corning 3881) at 80 μL per well. Eriochrome blue black B and ferriprotoporphyrin IX were suspended in 100% DMSO at 5 mM and then diluted in the peptide solution to reach a final DMSO concentration that did not exceed 1%.

ThT Assays. Assays were initiated by placing the 96-well plate at 37 °C under quiescent conditions in a plate reader (Fluostar Omega, Fluostar Optima or Fluostar Galaxy, BMGLabtech, Offenburg, Germany). The ThT fluorescence was measured through the bottom of the plate with a 440 nm excitation filter and a 480 nm emission filter. The ThT fluorescence was followed for three repeats of each sample for 15 h. Subsequently, the ThT absorbance at the end of the experiment was measured to determine the difference in the amount of Aβ42 fibrils generated in the presence of the small molecules compared to the control.

ASSOCIATED CONTENT

Supporting Information
The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscombsci.5b00129.

Compounds and libraries of compounds for Aβ, α-synuclein, and tau (PDF)

AUTHOR INFORMATION

Corresponding Author
*E-mail: mv245j@cam.ac.uk.

Notes
The authors declare no competing financial interest.

REFERENCES

(2) Dobson, C. M. Protein folding and misfolding. Nature 2003, 426, 884.


**NOTE ADDED AFTER ASAP PUBLICATION**

This article published with an incorrect version of the Supporting Information. The article reposted March 1, 2016 with the correct version.