# Small molecule sequestration of amyloid-β as a drug discovery strategy for Alzheimer's disease

Gabriella T. Heller[1], Francesco A. Aprile[1†], Thomas C. T. Michaels[1,2], Ryan Limbocker[1‡], Michele Perni[1], Francesco Simone Ruggeri[1], Benedetta Mannini[1], Thomas Löhr[1], Massimiliano Bonomi[3], Carlo Camilloni[4], Alfonso De Simone[5], Isabella C. Felli[6,7], Roberta Pierattelli[6,7], Tuomas P. J. Knowles[1], Christopher M. Dobson[1], Michele Vendruscolo[1*]

*[1]Centre for Misfolding Diseases, Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK*

*[2]Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge MA, USA*

*[3]Structural Bioinformatics Unit, Department of Structural Biology and Chemistry; CNRS UMR 3528; C3BI, CNRS USR 3756; Institut Pasteur, Paris, France*

*[4]Dipartimento di Bioscienze, Università degli Studi di Milano, 20133 Milano, Italy*

*[5]Division of Molecular Biosciences, Imperial College London, London SW7 2AZ, UK and Department of Pharmacy, University of Naples "Federico II", Naples, 80131 Italy*

*[6]Magnetic Resonance Center (CERM), University of Florence, 50019 Sesto Fiorentino, Italy*

*[7]Department of Chemistry "Ugo Schiff", University of Florence, 50019 Sesto Fiorentino, Italy*

*[†]Current address: Department of Chemistry, Molecular Science Research Hub, Imperial College London, London W12 0BZ, UK*

*[‡] Current address: Department of Chemistry and Life Science, United States Military Academy, West Point, NY 10996, USA*

**Corresponding author. Email: mv245@cam.ac.uk*

## Abstract

Disordered proteins are challenging therapeutic targets, and no drug is currently in clinical use that has been shown to modify the properties of their monomeric states. Here, we identify a small molecule, called 10074-G5, capable of binding and sequestering the intrinsically disordered amyloid-β peptide (Aβ) in its monomeric, soluble state. Our analysis reveals that this compound interacts with Aβ and inhibits both the primary and secondary nucleation pathways in its aggregation process. We characterise this interaction using biophysical experiments and integrative structural ensemble determination methods. We observe that this small molecule has the remarkable effect of increasing the conformational entropy of monomeric Aβ while decreasing its hydrophobic surface area. These results provide an illustration of the strategy of targeting the monomeric states of disordered proteins with small molecules to alter their behaviour for therapeutic purposes.

## Teaser

A small molecule binds to a disordered protein in its monomeric form, preventing its aggregation linked to Alzheimer's disease.

## Introduction

Alzheimer's disease is a fatal neurodegenerative condition that affects over 50 million people worldwide, a number that is predicted to rise to 150 million by 2050 unless methods of prevention or treatment are found (*1*). Despite over 25 years of intensive research and hundreds of clinical trials, there is still no drug capable of modifying the course of this disease (*1*).

The aggregation of the amyloid-β peptide (Aβ) in brain tissue is one of the hallmarks of Alzheimer's disease (*2-4*). This process involves at least three forms of Aβ: i) a monomeric state, which is highly disordered, ii) oligomeric aggregates, which are heterogeneous, transient and cytotoxic, and iii) fibrillar structures, which are ordered and relatively inert, although they are capable of catalysing the formation of Aβ oligomers (*3*). More generally, the aggregation of Aβ involves a complex non-linear network of inter-dependent microscopic processes, including: 1) primary nucleation, in which oligomers form from monomeric species, 2) elongation, in which oligomers and fibrils increase in size by monomer addition, 3) secondary nucleation, whereby the surfaces of fibrillar aggregates catalyse the formation of new oligomeric species, and 4) fragmentation, in which fibrils break into smaller pieces, increasing the total number of oligomers and fibrils capable of elongation (*2*).

Aβ is produced by proteolysis from the transmembrane amyloid precursor protein, and its 42-residue form (Aβ42) is the predominant species in deposits characteristically observed in the brains of patients with Alzheimer's disease (*2-4*). Kinetic analysis shows that, once a critical concentration of Aβ42 fibrils has been formed, secondary nucleation overtakes primary nucleation in becoming the major source of Aβ42 oligomers, as fibril surfaces act as catalytic sites for their formation (*3*). The fact that the oligomers appear to be the most toxic species formed during the aggregation process (*5, 6*), however, suggests that therapeutic strategies targeting Aβ aggregation should not primarily aim at inhibiting fibril formation per se, but rather doing so in a manner that specifically reduces the generation of oligomeric species (*2*). Complex feedback mechanisms between the different microscopic steps in the aggregation reaction can lead to an increase in the concentration of oligomers even when the formation of fibrils is inhibited, and hence result in an increase in pathogenicity (*2*).

Previous studies have suggested that effective strategies for inhibiting Aβ aggregation could be based on targeting fibril surfaces to supress the generation of oligomers, or on the reduction of

the toxicity of the oligomers (*7-11*). It is unclear, however, whether sequestering Aβ in its soluble state could be an effective drug discovery strategy against Alzheimer's disease. Stabilisation of monomeric Aβ into a β-hairpin conformation with large biomolecules has been previously demonstrated to inhibit aggregation, for example using an affibody protein (*12*). However, whether such stabilisation of Aβ in its monomeric form can be achieved via small molecule binding in a drug-like manner is still under debate. While there is research indicating a stabilising effect of small molecules on the soluble state of Aβ, there are contradictory reports of their effects on its aggregation (*13-15*). It should also be considered that such molecules may not be specific, as for example some appear to bind monomeric Aβ in a manner similar to low concentrations of sodium dodecyl sulphate (SDS) (*13-15*). Furthermore, it has been proposed that the binding of these small molecules to monomeric Aβ may be mediated by colloidal particles formed by the small molecules (*16*), although this observation has also been disputed (*13, 14*). The uncertainty of whether monomeric Aβ is a viable drug target is caused, in part, by a lack of understanding of the molecular properties of monomeric Aβ and how to stabilise this peptide with specific small molecules that have the potential to be developed as drugs.

The complexity of targeting monomeric Aβ is caused, in part, by the fact that Aβ is intrinsically disordered, as it lacks a well-defined structure and instead exists as a heterogeneous ensemble of conformationally distinct states (*17*). The dynamic nature of disordered proteins, and the consequent absence of stable and persistent binding pockets, implies that they do not readily lend themselves to conventional mechanisms of drug-binding, such as the well-established lock-and-key paradigm, in which a drug fits snugly into a single, well-defined binding site (*18, 19*). As a result, targeting disordered proteins with small molecules has not been considered a promising drug discovery strategy, and there are no small molecules on the market directly targeting disordered regions despite their high prevalence in disease. A deeper understanding of the possible mechanisms by which small molecules can modify the behaviour of disordered proteins may open new avenues for drug development, not only against Alzheimer's disease and other neurodegenerative disorders but also many other medical conditions involving disordered proteins, including type II diabetes, and certain forms of cancer and cardiovascular disease (*17, 18*). This insight may also be particularly valuable in modulating liquid-liquid phase separation, which often involves protein disorder (*20*).

Using experimental and computational biophysical techniques and mathematical modelling, we characterise the interaction of the small molecule 10074-G5 (biphenyl-2-yl-(7-nitro-

benzo[1,2,5]oxadiazol-4-yl)-amine, **Figure 1a**), with Aβ42 in its disordered, monomeric state. 10074-G5 has been previously identified to inhibit c-Myc-Max heterodimerization (*21*) specifically by binding and stabilizing the intrinsically disordered c-Myc monomer (*22, 23*). Here, we also observe that 10074-G5 binds monomeric Aβ42, a disordered peptide unrelated to c-Myc. As a result of this interaction, 10074-G5 significantly delays both primary and secondary nucleation pathways in Aβ42 aggregation. We characterise this interaction using biophysical experiments and integrative structural ensemble determination techniques, and observe that Aβ42 remains disordered in the bound form with decreased hydrophobicity. Notably, we also observe that the conformational entropy of Aβ42 increases upon interacting with 10074-G5, suggesting that exploiting this phenomenon may be a potential therapeutic strategy for disordered proteins.

## Results

### Selection of the system

We selected the compound 10074-G5 as model system to understand whether and how a small molecule inhibits the aggregation of Aβ by binding the monomeric form of this peptide. We used this molecule as it has been reported to bind the oncogenic disordered protein c-Myc in its monomeric form, and it contains a nitrobenzofurazan moiety, which has been previously shown to inhibit the aggregation of Aβ (*24*).

### Characterisation of the binding of 10074-G5 to monomeric Aβ42

We characterised the binding of 10074-G5 with monomeric Aβ42 using a multidisciplinary approach based on experiments and integrative structural ensemble determination. First, we carried out bio-layer interferometry (BLI, see **Materials and Methods**) measurements to characterise this interaction in real-time. We immobilised N-terminally biotinylated monomeric Aβ42 on the surface of super streptavidin sensor tips (**Materials and Methods**) and exposed them to varying concentrations of the small molecule (**Figures 1b** and **S1**). We observed a concentration-dependent response, indicative of binding. We globally fit the binding curves to simple one-step association and dissociation equations, such that the fit constrains all curves to share single association ($k_{on}$) and dissociation ($k_{off}$) rates, with a global $R^2$ value of 0.98. From these fits, we determined $k_{on}$ to be 8.5 x $10^3$ ± 0.2 x $10^3$ $M^{-1}sec^{-1}$ and $k_{off}$ to be 4.7 ×

$10^{-2} \pm 2 \times 10^{-4}$ $sec^{-1}$, corresponding to a binding dissociation constant ($K_D$) of 6 μM. This affinity value is comparable to other small molecule interactions with disordered proteins (*22*).

We then investigated the binding of 10074-G5 and monomeric Aβ42 at the ensemble-averaged, single residue level. To do so, we performed 2D $H^{N-BEST}CON$ nuclear magnetic resonance (NMR) experiments (*25*) on uniformly $^{13}C$, $^{15}N$-labelled monomeric Aβ42 in the presence of 1- and 2-fold concentrations of 10074-G5. As monomeric Aβ42 is relatively stable in solution at low concentrations and temperatures, we examined the binding of 10074-G5 to monomeric Aβ42 under these conditions (20 μM of Aβ42 of at 5 °C). Minimal chemical shift perturbations were observed in the 2D $H^{N-BEST}CON$ spectra upon the addition of compound at a 2:1 ligand:protein ratio at 5 °C (**Figure S2a**), consistent with other reports of small-molecule binders of disordered proteins (*22, 26*), suggesting that Aβ42 remains disordered in the presence of 10074-G5. We then performed this experiment at 15 °C in the absence and presence of pre-saturation of the solvent. This experiment, which relies on heteronuclear direct detection with minimal perturbation of proton polarization, provides a valuable tool to study solvent exposed systems in which amide protons experience fast hydrogen exchange (*25*). In particular, the signals of amide nitrogen atoms become attenuated when their directly bound protons are in fast exchange with the solvent. After testing this on a well characterized protein (ubiquitin), we performed this experiment on Aβ42 (**Figure 1c**). In the absence of 10074-G5, we observed that several hydrophobic residues, particularly C-terminal residues are protected from solvent exchange (Leu17, Leu34, Val36, Ile41, Ala42, see $I/I_0$ values as shown in **Figure 1d**). Notably, in the presence of 10074-G5, we observed the quenching of several residues across the sequence of the monomeric Aβ42 peptide that were not quenched in the absence of 10074-G5 (Glu3, Phe4, Val12, His13, Ala21, Asp23, Val24, Leu34, Val36, see $I/I_0$ values as shown in **Figure 1d**), suggesting that some residues have increased solvent exposure in the presence of the small molecule. This change in the solvent exchange profile suggests that 10074-G5 interacts with monomeric Aβ42 in a manner that increases the solubility of at least some of the conformations within the monomeric structural ensemble (*18*).

To obtain further insight into the thermodynamic properties of this interaction, we quantified the heat changes upon 10074-G5 binding to Aβ40 using isothermal titration calorimetry methods (**Figure S2b, c)**. In these experiments, we used Aβ40 instead of Aβ42 because of the higher solubility of Aβ40; we have, however, shown that 10074-G5 has similar effects on the

aggregation of Aβ40 as on that of Aβ42 (**Figure S3a)**. The observation of minimal heat changes (**Figure S2b, c)** suggests that the interaction of 10074-G5 with monomeric Aβ has minimal enthalpy and is thus likely to be entropic, as found for the interactions of another small-molecule with a disordered peptide (*27*).

To obtain a structural description of how 10074-G5 affects the disordered structural ensemble of Aβ42, we employed metadynamic metainference, an integrative structural ensemble determination approach (*28, 29*) that combines all-atom molecular dynamics simulations with NMR chemical shift data to improve force field accuracy (see **Materials and Methods**, **Figures 2, 3,** and **S4-6**). These simulations reveal that Aβ42 remains disordered in the form bound to 10074-G5, retaining several ensemble-averaged properties. In particular, we noted that most average inter-residue contacts of the unbound peptide remain the same in the bound form (**Figure 2a**). Furthermore, the distributions (**Figure 2b**) and averages values (apo: 1.146 ± 0.002, holo: 1.116 ± 0.004 nm) of the radii of gyration of the bound and unbound forms of the peptide are also highly similar. The presence of 10074-G5 does, however, alter the conformational ensemble of Aβ42, promoting conformations with lower relative hydrophobic surface area (the fraction of accessible hydrophobic surface area with respect to the total accessible surface area, **Figure 2c**). To determine which residues became more exposed or protected in the presence of 10074-G5, we compared the residue-specific average solvent accessible surface areas per residue in the presence and absence of 10074-G5 (**Figure 2d**). Notably, we observe that $Lys_{34}$ and $Met_{35}$ become more protected in the presence of the small molecule. $Met_{35}$ has previously been identified as a key residue for attention of aggregation, as oxidation has been shown to reduce the lag time of primary nucleation (*30*). These data represent distributions and averages of the conformational ensembles. While **Figure 1d** experimentally quantifies residue-specific changes in solvent exposure, the result is both ensemble- and time-averaged over the duration of the measurement and cannot be quantitively compared to residue-specific average solvent accessible surface areas calculated from these simulations (**Figure 2d**), as these results are not time-averaged.

The decrease in the hydrophobic surface area of Aβ42 prompted us to investigate the role of water in the binding interaction. To this end, we calculated the average number of hydrogen bonds formed by water molecules within the hydration shell (all waters within 0.4 nm of Aβ42). We found that water molecules form 3.07509 ± 0.00004 hydrogen bonds on average in the

absence of 10074-G5, a number that increases in its presence to 3.0959 ± 0.0005, getting closer to the bulk-like values of 3.44801 ± 0.00002 and 3.44791 ± 0.00002, for apo and holo systems, respectively, under the conditions investigated here; errors are calculated comparing the first and second halves of the trajectories (**Figure S6a**) (*31*). As expected, we observed instead little difference between the average number of hydrogen bonds formed by water molecules in the bulk between the apo and holo simulations (**Figure S6b**). To determine whether or not this binding could be characterised by the release of water molecules upon association, we calculated the average number of water molecules in the hydration shell and show that this value is similar with and without association (**Figure S6c**).

Given the lack of evidence for enthalpic contributions to the binding, we investigated whether or not we could detect entropic signatures of binding. Previously, changes in conformational entropy have been estimated from probability distributions of structural parameters, such as the radius of gyration (*32*). Given the similarity of the radius of gyration in the bound and unbound forms of the peptide (**Figure 2b**), to increase the sensitivity we applied this approach to the residue-specific Ramachandran plots (**Figure 3a**), finding that for the majority of non-terminal residues, the entropy increases in the bound form of the peptide. Nevertheless, quantifying the total change in the conformational entropy of Aβ42 in the presence and absence of 10074-G5 using this approach is challenging given the correlated nature of consecutive dihedral angles. To unequivocally quantify the differences in conformational entropy of the peptide upon binding the small molecule, structurally characterise the unbound and bound conformations of Aβ42, and assess convergence, we performed a clustering analysis (**Figures 3b-c** and **S6d**). We observed that while 10074-G5 binds the extended form in a non-specific manner, all other structural clusters show localisation of the compound within well-defined pockets of Aβ42 for specific conformations (**Figure 2c and 3b**) involving hydrophobic, hydrophilic, charged, and polar residues (**Figure S6d**). Interestingly, we also observed that the conformational entropy of Aβ42 is increased in the bound form, exhibiting the so-called entropic expansion mechanism, in which the binding of the protein contributes favourably to the binding free energy (*18*) (**Figure 3d**). Often, drug binding for folded proteins is described in terms of a lock-and-key binding mechanism, in which a protein is constrained by the binding of a small molecule, and therefore the entropy of the protein contributes unfavourably to the free energy of binding. In stark contrast to this, the entropic expansion binding mechanism suggests that the protein becomes more disordered upon interacting with a small molecule, and thus the conformational landscape is

expanded upon binding (*18*). This observation suggests that the identification of small molecules which increase the conformational entropy of the disordered proteins may be a promising therapeutic strategy.

To probe the energetic contributions to this interaction on an ensemble-averaged, residue-specific level, we analysed Lennard-Jones and Coulomb contributions between 10074-G5 and each residue. We observe strong Lennard-Jones interactions, particularly between aromatic residues (**Figure 2e**) $Tyr_{10}$, $Phe_{19}$ and $Phe_{20}$ and 10074-G5. The strongest Coulomb interactions occur at charged residues $Lys_{16}$ and $Lys_{28}$ (**Figure 2e**).

**The small molecule 10074-G5 sequesters monomeric Aβ42 and inhibits its aggregation**

We measured the kinetics of Aβ42 aggregation at a concentration of 1 μM in the presence and absence of increasing concentrations of 10074-G5. Measurements were performed by means of a fluorescent assay based on the amyloid-specific dye thioflavin T (ThT), which reports on the overall fibril mass formed during the aggregation process (*3, 33-36*). While the range of possible concentrations of Aβ42 and 10074-G5 were restricted for the NMR experiments due to a lack of sensitivity and limited solubilities of both Aβ42 and 10074-G5, the high sensitivity of ThT enabled us to work with lower concentrations of Aβ42 as compared to the NMR experiments and thereby probe higher ligand:protein ratios. We found that 10074-G5 has a significant effect on Aβ42 aggregation (**Figure 4a,b**). Specifically, the data show that the final value of the ThT fluorescence, which corresponds to the end point of the aggregation reaction, is dependent on the concentration of the compound (**Figure 4a**). The observation of a significant decrease in the final ThT intensity could be due to several non-mutually exclusive possibilities including: 1) interference of the ThT signal by 10074-G5, 2) formation of soluble off-pathway aggregates, 3) sequestration of Aβ42 during the aggregation process (*7*).

Given the fact that 10074-G5 is a coloured compound, we sought to investigate whether the decrease in the fluorescence intensity of ThT was exclusively due to an interference of 10074-G5 with the dye, or also due to a decrease in the mass of the fibrils formed during the aggregation process. To this end, we performed a ThT-independent dot-blot assay in which we explicitly measured the quantity of soluble Aβ42 over time in the presence and absence of 10074-G5 using the W0-2 antibody, which binds to Aβ (**Figure 4c-e**). The solubility was determined by measuring the amount of Aβ42 that did not sediment after 1 h of ultracentrifugation at 100,000 rpm. We observed that in the presence of a 20-fold excess of 10074-G5, approximately 40% of

the total amount of Aβ42 remained in a soluble form (**Figure 4d,e**). These experiments indicate that not all Aβ42 monomers are incorporated in ThT-binding fibrils at the end of the aggregation process, and, thus, that the presence of 10074-G5 sequesters Aβ42 in its soluble form. These dot-blot data can be explained by an equilibrium model of competitive binding, where monomers can bind both to amyloid fibril ends and to 10074-G5 (**Materials and Methods**, **Figure 4e**). A fit of the dot-blot data to this equilibrium model (Eq. S13), yields an affinity of 10074-G5 for the monomers of $K_D = 7 \pm 1$ μM, a value consistent with that determined independently from the BLI experiments ($K_D = 6$ μM), especially considering that in the BLI setup Aβ42 is constrained on a surface (**Figure 1c**) whereas the dot blot measurement is performed in bulk solution. We further confirmed the observation that Aβ42 remains soluble by exploiting the intrinsic fluorescence of Tyr10 in the Aβ42 sequence. By monitoring the aggregation of 5 μM Aβ42 from its monomeric form over 1 h, the fluorescence intensity of $Tyr_{10}$ increases considerably (**Figure 4f**) as it becomes buried in a hydrophobic environment in the aggregated state (*37*). This experiment reports on early aggregation events that may be invisible to ThT, which is specific for cross-β sheet content as early aggregates such as oligomers or multimers may lack β-sheet structure (*38*). We observed, however that in the presence of an equimolar concentration of 10074-G5, the fluorescence intensity remains constant over time (**Figure 4g**), thereby suggesting that Aβ42 does not self-associate in the presence of 10074-G5.

To further demonstrate that 10074-G5 alters the kinetics of aggregation, we performed 3-D morphological analyses of fibrils using high resolution and phase-controlled (*39*) atomic force microscopy (AFM) on the time scale of the aggregation process (**Figures 4b** and **S7**). Single-molecule statistical analysis of the aggregates in the 3-D maps shows that Aβ42, both in the absence and presence of 10074-G5, forms non-mature aggregates with average cross-sectional diameters of approximately 2-3 nm, and mature fibrillar aggregates with average diameters of approximately 5-6 nm, as previously observed (*40, 41*). It has been shown that fibrillar species with diameters less than 6 nm lack a complete (or mature) cross-β sheet structure stabilised by a tight network of intermolecular hydrogen bonding, as compared to mature fibrillar aggregates (*40*). Notably, we observed that at the same time point of aggregation, the fibrillar aggregates formed in the presence of 10074-G5 had smaller cross-sectional diameters than those formed in its absence, with a significantly higher abundance of non-mature species with respect to mature fibrillar species. These data suggest that the process of fibril formation and maturation of cross-

β sheet structure in the presence of this compound is considerably slower than in its absence (*41*) (**Figures 4b** and **S7**).

**10047-G5 does not chemically modify Aβ42**

To determine whether or not the binding of 10074-G5 to Aβ42 is covalent or induces other chemical modifications, we performed mass spectrometry on Aβ42 incubated in the presence and absence of 10074-G5. Samples were incubated overnight at 37 °C and then spun down using an ultracentrifuge (**Materials and Methods**). The supernatant and resuspended pellet of the aggregation reactions were analysed by matrix assisted laser desorption/ionization (MALDI) mass spectrometry (**Figure S8**). No mass increase was observed following the incubation with 10074-G5, indicating that its presence does not result in detectable covalent chemical modifications to Aβ42.

**10074-G5 inhibits all microscopic steps of Aβ42 aggregation**

In order to better understand the mechanism of inhibition of Aβ42 aggregation by 10074-G5, we performed a kinetic analysis on the ThT aggregation traces. **Figure 5a** shows the ThT kinetic curves, and from the data, we observe that 10074-G5 slows down the aggregation reaction in a concentration-dependent manner, consistent with the AFM results, confirming a delay in the aggregation process of Aβ42 (**Figure 4b** and **S7**). Furthermore, our dot blot analysis shows that a fraction of the total monomer concentration is left unreacted at the end of the aggregation reaction (**Figure 3c-e**). The simplest model that accounts both for the delay of aggregation and for the fact that not all Aβ ends up in the fibrils is one in which monomer is effectively removed from the reaction by the inhibitor. Indeed, the final load on fibrils depends on the amount of free monomer available for the reaction. Furthermore, reducing the pool of available monomers slows down the rates of all aggregation steps, as monomers are involved in all microscopic steps.

Models of inhibition in which fibril surfaces or fibril ends are targeted by inhibitors would explain the delay in aggregation, but not the decrease fibril load (**Figure S3d**). By using explicit rate laws (*36*) for describing this inhibition model by monomer sequestration, we sought to understand whether our data could be explained by 10074-G5 exclusively binding the monomer, and in this manner, reducing the concentration of monomers available for the overall aggregation reaction and each microscopic step of aggregation (see **SI**). Specifically, we first fitted the

measured aggregation kinetics in the absence of 10074-G5 to a kinetic model of Aβ42 aggregation (see **SI**, Eq. S10) (*36*) to estimate the values of the unperturbed rates for primary nucleation, elongation, and secondary nucleation. We then formulated a master equation model for inhibited aggregation kinetics in the presence of 10074-G5 (Eq. S11). We derived explicit integrated rate laws describing inhibited kinetics (Eqs. S11-14 and **Figure S9**), which we used to fit the experimental ThT data in the presence of 10074-G5. For this analysis, we implemented the unperturbed rate constants for aggregation, leaving the value of $K_D$ as the only fitting parameter. We performed a global fit to the normalized data, as described previously (*42*); all ThT profiles at increasing concentrations of 10074-G5 were not fit individually, but rather using the same choice of $K_D$, with the dependence of the various rate constants on the concentration of 10074-G5 being captured in the integrated rate law through Eq. S14. The result of this global fit, shown in **Figure 5a,** accounts for the retardation of aggregation in a global manner using one single parameter ($K_D$) and explains the observed decrease in final fibril load. Furthermore, this fit yields an affinity value of $K_D =$ 40 μM, consistent with other affinities reported for small molecule binders of disordered proteins (*22*). It is interesting to note that this estimated affinity is considerably weaker than small molecule binders of structured proteins, which are often in the nM range. It may seem that such a weak binder of Aβ would have little inhibitory effects. However, we note that the level of inhibition in general depends not on the absolute value of $K_D$, but on the combined parameter $K_D/[C]$, where [*C*] is the drug concentration, provided that the rate of binding ($k_{on}[C]$) is sufficiently fast as compared to the overall timescale of aggregation ($k$) (*43*). Therefore, it is possible to have effective inhibition even for small molecules with $K_D$ values in the μM range, provided that $[C]$ is on the same order of magnitude as $K_D$ (see **Figure 6**).

The analysis of experimental aggregation data in the presence of increasing concentrations of inhibitor using our integrated rate law thus yields an independent method for determining the binding constant of 10074-G5 to the monomers. To provide further support to this analysis, we varied the concentration of monomeric Aβ42 (1, 1.5 and 2 μM) and recorded kinetic traces of aggregation in the absence (**Figure 5b**) and presence of 10 μM 10074-G5 (**Figure 5c**). Using the rate parameters determined from the uninhibited kinetics and the same value of $K_D$ obtained from the global fit shown in **Figure 5a**, we find that the time course of aggregation predicted by our monomer sequestration model are in good agreement with the independent experimental data (**Figure 5c**).

A key prediction from the monomer sequestration model is that a monomer-interacting compound should interfere with all three microscopic steps of aggregation. In fact, we find that the presence of an inhibitor that binds monomers quickly compared to the overall aggregation rate does not affect the topology of the reaction network. As a result, the inhibited kinetics can be interpreted in terms of effective rates of aggregation that depend on the concentration of inhibitor (Eq. S14). In **Figure 5d**, we show the values of the effective rates of aggregate proliferation through primary ($\lambda$) and secondary ($\kappa$) nucleation pathways as a function of the concentration of 10074-G5 predicted by this model (see Eq. S10 for a definition of $\lambda$ and $\kappa$). The monomer sequestration model also predicts that the effective rate of elongation should be reduced, although to a lesser extent than the nucleation pathways, which have a stronger monomer concentration dependence. To test this prediction, we performed seeded aggregation experiments in the presence of preformed Aβ42 fibrils to obtain independent measurements of the effective elongation rate as a function of 10074-G5 concentration. We observed that 10074-G5 indeed decreases the effective rate of fibril elongation (**Figure S3b**), consistent with the monomer sequestration mechanism.

We also sought to understand whether 10074-G5 binds monomeric Aβ40 with a comparable affinity to that of Aβ42. To address this question, we applied the monomer sequestration model as used in **Figure 5a**, to fit the inhibitory effects of 10074-G5 on aggregation of 10 μM of Aβ40. From this analysis, we extracted an affinity constant of 10074-G5 for Aβ40 of 10 μM, a similar value to that obtained for Aβ42 (**Figure S3a**). Given the increased toxicity of Aβ42 as compared to Aβ40, we anticipate the optimisation of small molecules more specific for monomeric Aβ42 over Aβ40.

**Characterisation of the binding of 10074-G5 to stabilised Aβ40 oligomers**

Next, we probed whether 10074-G5 alters the behaviour of oligomeric species of Aβ. Although it is extremely challenging to determine whether 10074-G5 modifies the oligomeric intermediates of Aβ42 formed on-pathway to aggregation, which are transient, heterogenous species, it is possible to carry out this analysis more readily on oligomers of Aβ40 stabilised using $Zn^{2+}$ (*44*). Thus, we next considered whether or not 10074-G5 can alter the behaviour of these stabilised, pre-formed oligomeric species. We incubated pre-formed oligomers in the presence of 10074-G5, centrifuged the samples, and measured the quantities of Aβ40 in the pellet and in the supernatant by using SDS–polyacrylamide gel electrophoresis (SDS-PAGE, **Figure**

**S10a**). The results indicate that these pre-formed oligomers did not dissociate in the presence of 10074-G5. Furthermore, 10074-G5 was found not to alter the turbidity of solutions in which they were present as measured by their absorbance profiles (**Figure S10b**) suggesting that 10074-G5 does not cause such species to change detectably in size. Lastly, dot blots of pre-formed oligomeric samples in the presence and absence of the compound using the OC-antibody, which binds to β-sheets (*45*), show that the oligomers maintain their characteristic conformations (**Figure S10c**). Due to the coloured nature of 10074-G5, it was neither possible to characterise the oligomers in the presence of the compound with dynamic light scattering nor analytical ultracentrifugation measurements. Taken together, these data suggest that 10074-G5 does not disaggregate the pre-formed oligomeric species or cause them to undergo further assembly. Nevertheless, it remains possible that this compound affects the evolution of oligomer populations formed during the aggregation reaction, potentially inhibiting their conversion into fibril-competent species.

## Conclusions

We have characterised the binding of the small molecule 10074-G5 to monomeric Aβ42 using a combination of experimental approaches and integrative structural ensemble determination methods. The real-time, dose-dependent responses that we have observed in the BLI experiments demonstrate that 10074-G5 binds Aβ42 in its monomeric form (**Figures 1b and S1**). The NMR experiments and the metadynamic metainference simulations have illustrated that this binding is distinct from most small molecule interactions with structured proteins. In particular, we have observed that 10074-G5 does not bind to a single binding site, but rather binds transiently to many different sites (**Figures 2, 3,** and **S6**). We have also found 10074-G5 induces Aβ42 to adopt many different conformations, keeping it disordered and more solvent-exposed (**Figure 1**) in the bound form. It appears that this interaction is largely entropic (**Figure S2b**) and that some of this entropy arises from the increase in conformational entropy of Aβ42 (**Figure 3**) via the entropic expansion mechanism (*18*). This structural insight may open new routes for inhibitors of pathogenic disordered proteins and small molecule modulators of liquid-liquid phase separation involving disordered regions. In particular, a greater understanding of how one could identify small molecules that alter conformational entropy of disordered proteins without extensive off-target effects holds enormous therapeutic potential.

Additionally, we have characterised the effects of 10074-G5 on amyloid aggregation *in vitro* using a range of biophysical techniques and kinetic theory techniques. This analysis has revealed that as a result of the interaction of 10074-G5 with monomeric Aβ42, this small molecule also reduces the extent to which monomeric Aβ42 contributes to aggregation, thereby effectively slowing down all microscopic aggregation rates. Our kinetic analysis of aggregation inhibition by monomer sequestration highlights that effective inhibition depends on the interplay of two combined parameters, namely a thermodynamic and kinetic one: $K_D/[C]$ and $k_{on}[C]/\kappa$ (**Figure 6**) (*43*). Thus, for a small molecule with given affinity for monomeric Aβ42, increasing $k_{on}$ of small molecules for monomeric Aβ42 represents a promising inhibitor optimisation strategy. We also anticipate future work to understand the relationship between $k_{on}$ and the change in conformational entropy of Aβ42 towards improved binders of monomeric disordered proteins.

Collectively, these results indicate the importance of developing a more detailed understanding of the interactions between disordered proteins and small molecules, which in turn could lead to the development of new therapeutic approaches for the many diseases in which such disordered proteins are involved.

**Acknowledgements**

GTH is supported by the Gates Cambridge Trust and the Rosalind Franklin Research Fellowship at Newnham College, Cambridge, FAA by UK Research and Innovation (MR/S033947/1) and the Alzheimer's Society UK (317, 511), RL by the Gates Cambridge Trust, TCTM by Peterhouse, Cambridge and the Swiss National Science Foundation, and FSR by Darwin College and the Swiss National Foundation (Grant Numbers P300P2_171219 and P2ELP2_162116, respectively). We acknowledge ARCHER UK National Supercomputing Service under ARCHER Leadership project (Grant Number e510) and PRACE for awarding us access to MareNostrum at Barcelona Supercomputing Center (BSC), Spain for metadynamic metainference simulations. Parameterisation of 10074-G5 was performed using resources provided by the Cambridge Service for Data Driven Discovery (CSD3) operated by the University of Cambridge Research Computing Service (www.csd3.cam.ac.uk), provided by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (capital grant EP/P020259/1), and DiRAC funding from the Science and Technology

Facilities Council (www.dirac.ac.uk). MALDI mass spectrometry measurements were performed by Dr. Len Packman at the Protein and Nucleic Acid Chemistry Facility (PNAC) at the Department of Biochemistry, University of Cambridge. The NMR measurements were supported by the iNEXT H2020 Programme (EC contract n. 653706). The work was also supported by the Centre for Misfolding Diseases and the INCEPTION project ANR-16-CONV-0005.

**Data Availability Statement**

The data and code that support the findings of this study are available from GitHub at https://github.com/vendruscolo-lab/amyloid-beta_small_mol.

All the data and PLUMED input files required to reproduce the metadynamic metainference results reported in this paper are available on PLUMED-NEST (www.plumed-nest.org), the public repository of the PLUMED consortium(*46*), as plumID:20.014.

**Competing Interests**

All authors declare that they have no competing interests.

**Author Contributions**

GTH, FAA, and MV designed the research. GTH and FAA performed BLI experiments, ThT aggregation assays, dot blots, ITC experiments, and fluorescence quenching measurements. GTH, ICF, and RP performed NMR experiments. RL and FSR performed AFM experiments. MP and RL performed C. elegans experiments. BM and GTH performed oligomer experiments. TCTM and TPJ performed kinetic modelling of ThT aggregation data. GTH, MB, TL, CC, and ADS performed metadynamic metainference simulations GTH, FAA, TCTM, RL,MP, FSR,BM,TL,MB,CC, ADS,ICF, RP, TPJK, CMD, and MV wrote the paper.

**List of supplementary materials**

The supplementary materials document contains details of the metadynamic metainference simulations, details of the kinetic analysis of experimental aggregation data using a monomer sequestration model, and 10 supplementary figures.

## Materials and Methods

*BLI experiments.* A super streptavidin biosensor (FortéBio, Menlo Park, USA) was coated with 15 μg/ml monomeric N-terminally biotinylated Aβ42 (AnaSpec, Fremont, USA) by overnight incubation a solution at 5 °C. Control biosensors were incubated with the same concentration of biocytin. The tips were then rinsed by incubation in buffer for 3 h at room temperature. The binding and dissociation between immobilized Aβ42 and various concentrations of 10074-G5 was monitored for 200 s and 500 s respectively at 37 °C using an Octet Red96 (ForteBio, Menlo Park, USA). This process was repeated six times (**Figure S1a, b**). The binding of buffer to a Aβ42-functionlized biosensor was subtracted to account for baseline drift. Data were analysed using GraphPad Prism 8. Dissociation data were first globally fit using a one-phase exponential decay to determine a preliminary $k_{\mathrm{off}}$ value. This value was then used as an initial input value to determine the global $k_{\mathrm{on}}$ and $k_{\mathrm{off}}$ rates.

*2D $H^{N-BEST}$CON NMR experiments.* $^{13}$C, $^{15}$N uniformly labelled, recombinant Aβ42 peptide (the 42-residue variant lacking the N-terminal M, see '*Preparation of recombinant Aβ peptides*') was purchased from rPeptide and prepared following the manufacturer's instructions. 20 μM samples were prepared in PBS (pH 7.50), 1% DMSO, with 5% $D_2O$ (Sigma Aldrich) for the lock. 2D $H^{N-BEST}$CON measurements (*25*) were performed at 16.4 T on a Bruker Avance spectrometer operating at 700.06 MHz $^{1}$H, 176.03 MHz $^{13}$C and 70.9 MHz $^{15}$N frequencies, equipped with a triple-resonance cryogenically cooled probehead optimized for $^{13}$C-direct detection (at the Centro di Risonanze Magnetiche, Florence, Italy). Each 2D $H^{N-BEST}$CON spectrum was acquired with 64 scans. The dimensions of the acquired data were 1024 ($^{13}$C) x 116 ($^{15}$N) points. The spectral width was 29.9 x 33.9 ppm for $F_2$ and $F_1$, respectively. The relaxation delay was set to 0.3 s. 2D $H^{N-BEST}$CON measurements were repeated with the same parameters except for the inclusion of a weak pre-saturation of the solvent signal during the relaxation delay. Under these conditions, signals of amide nitrogen whose directly bound protons are in fast exchange with the solvent are attenuated. This approach was tested on a well characterized protein (ubiquitin) and then used for the study of the Aβ42 peptide with and without the addition of 10074-G5. 1D $^{1}$H and 2D BEST TROSY (*47*) spectra were acquired before and after measurements were taken to ensure that minimal aggregation had occurred during the course of the measurement. Experimental data were acquired at 5 and 15 °C using Bruker TopSpin 3.1 software and processed with Sparky 3.115.

*Metadynamic metainference simulations.* To generate the structural ensembles, we employed an integrative approach that incorporates NMR chemical shift data into molecular dynamics simulations. To this end, we used metadynamic metainference, which compensates for the inaccuracies of the force field, accounts for errors in experimental data, and enhances sampling (*28, 29*). All-atom metadynamic metainference simulations (*28*) of the unbound and bound form of Aβ42 were performed using GROMACS 2018.3 (*48*) patched with PLUMED library (*46*) 2.6.0-dev (git: 0edcfb268569) (*46*), the CHARMM22* force field (*49*), and TIP3P water model (*50*). The initial conformation of Aβ42 was prepared as a linear peptide using PyMol (https://pymol.org/2/). A preliminary *in vacuo* molecular dynamics simulation was performed for 1 ns to collapse the extended conformation. This structure was solvated in a rhombic dodecahedron box with an initial volume of 362 $nm^3$ containing 11,746 water molecules. The solvated system was minimised using the steepest descent algorithm with a target maximum force of 1,000 kJ $mol^{-1}$ $nm^{-1}$. A pool of 48 initial conformations was extracted from a preliminary 2 ns simulation at 600 K in the NVT ensemble. Equilibration was then performed in the NVT ensemble for 500 ps at 278 K using the Bussi-Donadio-Parrinello thermostat (*51*) and for 500 ps at 278 K in the NPT ensemble using Berendsen pressure coupling (*52*) with position restraints on heavy atoms. Production runs were executed in the NPT ensemble at 278 K using the Parrinello-Rahman barostat (*53*). A time step of 2 fs was used together with LINCS constraints on all bonds (*54*). The van der Waals interactions were cut-off at 1.2 nm, and the particle-mesh Ewald method was used for electrostatic interactions (*55*). Bound simulations were performed as described above, using the starting structures obtained from the NVT equilibration at 600 K. The 10074-G5 molecule was added to a corner of the box and the system re-solvated with 11,734 water molecules. The system was then minimized and equilibrated using the procedures described above. Preliminary parameters for 10074-G5 were taken from the CGenFF software (*56*), and those with any penalty were explicitly re-parameterised using the Force Field Toolkit (*57*) and Gaussian 09 (http://www.gaussian.com) (see **SI** and **Figure S4**).

Chemical shifts were back-calculated at each time step using CamShift (*58*) (**Figure S5**). Given that the error of the CamShift predictor is greater than the chemical shift perturbations upon addition of the compound, the same chemical shifts were used to restrain both the unbound and bound simulations. A Gaussian noise model with one error parameter per nucleus type was used in the metainference setup, along with an uninformative Jeffreys prior for each error parameter

(*28*) (see **SI**). The metainference ensembles for the unbound and bound simulations were simulated using 48 replicas each.

Parallel bias metadynamics (*59*) with the well-tempered (*60*) and multiple-walkers (*61*) protocols was performed using a Gaussian deposition stride of 1 ps, an initial height of 1.2 kJ/mol, and bias factors of 24 and 49 for the unbound and bound simulations, respectively. In the unbound simulations, we used 6 collective variables (CVs) to enhance the conformational sampling of Aβ42 (see **SI**). In the bound simulations, we also included 14 CVs to enhance the conformational sampling of contacts between the compound and the peptide, and 4 CVs to enhance sampling of soft dihedrals in the small molecule (see **SI**). Unbound and bound simulations were run for an accumulated time of 27.8 and 28.2 μs, respectively until convergence was reached (see **SI** and **Figure 3c**). For details on the analysis, see **SI**.

*Preparation of recombinant Aβ peptides.* Recombinant Aβ(M1-42) (MDAEFRHDSGY EVHHQKLVFF AEDVGSNKGA IIGLMVGGVVIA) and Aβ(M1-40) (MDAEFRHDSGY EVHHQKLVFF AEDVGSNKGA IIGLMVGGVV), here referred to as Aβ42 and Aβ40, respectively, were prepared by expression in *Escherichia coli* BL21 (DE3) Gold Strain (Agilent Technologies, Santa Clara, USA) (*35*). The resulting inclusion bodies were dissolved in 8M urea, ion exchanged in batch mode on diethylaminoethyl cellulose resin, lyophilized, and then further purified with a Superdex 75 HR 26/60 column (GE Healthcare, Chicago, USA). Those fractions containing the recombinant protein, as determined by SDS–polyacrylamide gel electrophoresis, were combined and lyophilized again. To ensure we were working with highly purified monomeric species containing extremely low quantities of aggregated forms of the peptides, size exclusion chromatography was performed directly before the experiments were performed. Aβ40 and Aβ42 solutions were prepared by dissolving the lyophilized peptide in 6 M GuHCl and incubating on ice for 3 h. The solutions were then purified using a Superdex 75 Increase 10/300 GL column (GE Healthcare, Chicago, USA) at a flow rate of 0.5 ml/min and eluted in 20 mM sodium phosphate buffer (pH 8) supplemented with 200 μM EDTA. The center of the peak was collected and the concentrations of the peptides were determined from the integration of the absorbance peak using $\varepsilon_{280}$=1,495 liter $mol^{-1}$ $cm^{-1}$.

*Preparation of small molecules.* 10074-G5 was obtained from Sigma Aldrich (St. Louis, MO, USA). The molecules were dissolved in 100% DMSO and then diluted in solutions of Aβ40 or

Aβ42 to reach a maximum final DMSO concentration of 1.5%. The total DMSO concentration was matched in the control solutions in all experiments.

*ThT aggregation kinetics.* Monomeric Aβ40 or Aβ42 were diluted with buffer and 20 μM ThT from a 2 mM stock and increasing amounts of 10074-G5. Samples were prepared using LoBind Eppendorf tubes (Sigma Aldrich, MO, St. Louis, USA) on ice. Fibrils for seeding experiments were prepared by incubating monomeric Aβ42 at 37 °C overnight. The concentration of fibrils (in monomer equivalents) was assumed to be the initial concentration of monomer. These preformed fibrils were added to a freshly prepared monomer solution to give a final concentration of 15% fibrils.

Samples with or without seed fibrils were pipetted into multiple wells of a 96-well half-area, low-binding polyethylene glycol coating plate (Corning 3881, Corning, USA) with a clear bottom, at 90 μl per well. Plates were sealed with aluminium sealing tape (Corning) to prevent evaporation and then placed at 37 °C under quiescent conditions in a plate reader (CLARIOstar; BMG Labtech, Ortenberg, Germany). ThT fluorescence was measured through the bottom of the plate using 440-nm and 480-nm excitation and emission filters, respectively. ThT fluorescence was followed in with multiple replicates as specified in the corresponding figure captions. For analysis of ThT kinetics see **SI**.

*Mass spectrometry.* Monomeric Aβ42 was diluted in the aggregation buffer (described above) to a concentration of 15 μM in the presence and absence of 30 μM 10074-G5. Samples were incubated overnight at 37 °C under quiescent conditions to mimic the aggregation experiments. The samples were then spun down using an ultracentrifuge at 100,000 rpm for 1 h at 25 °C to separate the supernatant and pellet. 6 M GuHCl was used to dilute the supernatant by 50% with and resuspend the pellet. Samples were analysed by MALDI mass spectrometry at the Protein and Nucleic Acid Chemistry Facility (PNAC) at the Department of Biochemistry, University of Cambridge.

*Dot-blot assay.* Blotting was performed using the Aβ42 sequence-specific antibody (W0-2, MABN10, Millipore, Burlington, USA). Samples were removed from a solution containing 2 μM Aβ42 in the presence and absence of three- and ten-fold equivalence of 10074-G5. To ensure only the monomer was placed on the blots, samples were spun down using an ultracentrifuge at

100,000 rpm for 1 h at 25 °C using a TLA100 rotor. 2 μL of the supernatant were pipetted onto a nitrocellulose membrane (0.2 μM; Whatman). After drying, the membrane was blocked with 5% (w/v) bovine serum albumin (BSA) in phosphate-buffered saline (8 mM $Na_2HPO_4$, 15 mM $KH_2PO_4$, 137 mM NaCl, 3 mM KCl, pH 7.4, PBS) overnight at 5 °C, followed by three 15 min washes with PBS at room temperature. The membrane was then immunised with a 1/1000 dilution of WO-2 anti-Aβ antibody in PBS with 5% BSA overnight at 5 °C, followed by three 15 min washes with PBS at room temperature. The membrane was then incubated for 2 h at room temperature in PBS supplemented with 0.05% tween and an anti-mouse-Alexa Fluor 594 secondary antibody conjugate (ThermoFisher Scientific, Waltham, USA) at room temperature, and then washed three times with PBS supplemented with 0.05% tween. Fluorescence detection was performed using Typhoon Trio Imager (GE Healthcare, Chicago, IL, USA). Blots were quantified using ImageJ. Data were fit to a competitive binding equilibrium model between free monomers and fibrils (Eq. S13). In this model, monomers are either free, aggregated (i.e. part of a fibril), or bound to 10074-G5; the binding of the compound to the monomer is described by a single binding free energy. The binding of monomers to fibril ends is stronger compared to the binding of monomers to the inhibitor. The concentration of free monomer in equilibrium with amyloid fibrils (critical concentration) measured in our experiments is $m_{critical}$ = 93 nM, consistent with other reports (*62*). The equilibrium concentration of unreacted soluble monomer after ultracentrifugation measured at varying inhibitor concentration is fit to Eq. S13 with $K_D$ as a fitting parameter. This procedure yields $K_D = 7 \pm 1$ μM, as shown in **Figure 4e**.

*Atomic force microscopy.* Solutions of 1 μM Aβ42 in the presence and absence of 6 μM 10074-G5 were deposited on mica positively functionalized with (3-aminopropyl) triethoxysilane (APTES, Sigma Aldrich, St. Louis, USA) in the absence of ThT. The incubation times were selected based on the results of the chemical kinetics experiments. The mica substrate was positively functionalized by incubation with a 10 μl drop of 0.05% (v/v) APTES in Milli-Q water for 1 min at ambient temperature, rinsed with Milli-Q water and then dried by the passage of a gentle flow of gaseous nitrogen (*40*). AFM sample preparation was carried out at room temperature by deposition of a 10 μL drop of protein solution deposited for 2 min to a surface treated with APTES. The samples were rinsed with Milli-Q water, dried with nitrogen gas, and stored in a sealed container until imaging. AFM maps were acquired by means of a NX10 (Park Systems, Suwon, Korea) and a nanowizard2 (JPK Instruments, Berlin, Germany) system operating in tapping mode and equipped with a silicon tip (PPP-NCHR and μmasch) with a

nominal radius of 10 nm. Image flattening and single aggregate statistical analysis were performed by SPIP 6 (Image Metrology, Hørsholm, Denmark) software.

*ITC experiments.* Isothermal titration calorimetry (ITC). Measurements were performed using an MicroCal Auto-ITC 200 (GE Healthcare, Chicago, USA) at 15°C. Due to the poor solubility of 10074-G5, monomeric Aβ40 (200 μM) was injected 10 times into a solution containing 7 μM of 10074-G5. All solutions were prepared in phosphate buffer (described above) and contained a minimal amount of dimethyl sulfoxide (DMSO, 0.2%) to ensure that the compound was soluble. Each injection was 3.5 μL in volume and was made on 3 min intervals. Heats of dilution, obtained by separately injecting the peptide into buffer and buffer into the solution containing 10074-G5, were subtracted from the final data. The corrected heats were divided by the number of moles injected and analyzed using Origin 7.0 software (OriginLab, Northampton, USA).

*Characterization of the interaction of 10074-G5 with stabilized oligomers.* Stabilised oligomers were formed from Aβ40 as previously described (*44*). Briefly, 1 mg of lyophilized Aβ40 was dissolved in 300 μL of hexafluoroisopropanol and incubated overnight at 4 °C. After solvent evaporation under nitrogen gas, Aβ40 was resuspended in DMSO to a concentration of 2.2 mM and sonicated twice for 10 min at room temperature. The protein sample was diluted to a final concentration of 100 μM in 20 mM sodium phosphate buffer with 200 μM $ZnCl_2$ at pH 6.9. After incubation for 20 h at 20 °C, the solution was centrifuged for 15 min at 15,000 *g* at room temperature. The pellet containing the oligomers was resuspended in 20 mM phosphate buffer at pH 6.9, with 200 μM $ZnCl_2$.

Samples containing 20 μM and 10 μM pre-formed $Zn^{2+}$-stabilised Aβ40 oligomers were incubated in the presence and absence of 20 μM 10074-G5 for 1 h. The turbidimetries of the samples were analysed using a plate reader (BMG Labtech, Aylesbury, UK) at 600 nm. Measurements were background subtracted against buffer alone in the absence and presence of compound. The protein content within samples was quantified using the sequence-specific WO-2 antibody (see *Dot-blot assay*). Similarly, the conformations of the oligomers in the presence and the absence of the compound was probed using the conformation-specific OC antibody (*45*) (AB2286, Millipore, Burlington, USA) using the protocols described above (see *Dot-blot assay).*

To determine if the oligomers had dissociated after the incubation in the presence of the compound, the samples were centrifuged for 15 min at 15,000 *g*. The pellet was resuspended in 15 μL of 20 mM phosphate buffer at pH 6.9 with 200 μM $ZnCl_2$ and analysed along with the supernatant by SDS–polyacrylamide gel electrophoresis.

**References**


1. J. Cummings, G. Lee, A. Ritter, K. Zhong, Alzheimer's disease drug development pipeline: 2019. *Alzheimer's Dementia* **5**, 272-293 (2019).
2. T. P. Knowles, M. Vendruscolo, C. M. Dobson, The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **15**, 384 (2014).
3. S. I. Cohen *et al.*, Proliferation of amyloid-β42 aggregates occurs through a secondary nucleation mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 9758-9763 (2013).
4. J. Hardy, D. J. Selkoe, The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* **297**, 353-356 (2002).
5. C. Haass, D. J. Selkoe, Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid β-peptide. *Nat. Rev. Mol. Cell Biol.* **8**, 101 (2007).
6. I. Benilova, E. Karran, B. De Strooper, The toxic Aβ oligomer and Alzheimer's disease: an emperor in need of clothes. *Nat. Neurosci.* **15**, 349 (2012).
7. P. Arosio *et al.*, Kinetic analysis reveals the diversity of microscopic mechanisms through which molecular chaperones suppress amyloid formation. *Nat. Commun.* **7**, 10948 (2016).
8. J. Habchi *et al.*, Systematic development of small molecules to inhibit specific microscopic steps of Aβ42 aggregation in Alzheimer's disease. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E200-E208 (2017).
9. J. Habchi *et al.*, An anticancer drug suppresses the primary nucleation reaction that initiates the production of the toxic Aβ42 aggregates linked with Alzheimer's disease. *Sci. Adv.* **2**, e1501244 (2016).
10. F. A. Aprile *et al.*, Selective targeting of primary and secondary nucleation pathways in Aβ42 aggregation using a rational antibody scanning method. *Sci. Adv.* **3**, e1700488 (2017).
11. S. I. Cohen *et al.*, A molecular chaperone breaks the catalytic cycle that generates toxic Aβ oligomers. *Nat. Struct. Mol. Biol.* **22**, 207 (2015).

12. W. Hoyer, C. Grönwall, A. Jonsson, S. Ståhl, T. Härd, Stabilization of a β-hairpin in monomeric Alzheimer's amyloid-β peptide inhibits amyloid formation. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 5099-5104 (2008).
13. C. Lendel, B. Bolognesi, A. Wahlström, C. M. Dobson, A. Gräslund, Detergent-like interaction of Congo red with the amyloid β peptide. *Biochemistry* **49**, 1358-1360 (2010).
14. A. Abelein, B. Bolognesi, C. M. Dobson, A. Gräslund, C. Lendel, Hydrophobicity and conformational change as mechanistic determinants for nonspecific modulators of amyloid β self-assembly. *Biochemistry* **51**, 126-137 (2011).
15. P. Pratim Bose *et al.*, Effects of congo red on Aβ1− 40 fibril formation process and morphology. *ACS Chem. Neurosci.* **1**, 315-324 (2010).
16. B. Y. Feng *et al.*, Small-molecule aggregates inhibit amyloid polymerization. *Nat. Chem. Biol.* **4**, 197 (2008).
17. J. Habchi, P. Tompa, S. Longhi, V. N. Uversky, Introducing protein intrinsic disorder. *Chem. Rev.* **114**, 6561-6588 (2014).
18. G. T. Heller, P. Sormanni, M. Vendruscolo, Targeting disordered proteins with small molecules using entropy. *Trends Biochem. Sci.* **40**, 491-496 (2015).
19. G. T. Heller, M. Bonomi, M. Vendruscolo, A structural ensemble modulation mechanism of small molecule binding to disordered proteins. *J. Mol. Biol.*, (2018).
20. V. Csizmok, A. V. Follis, R. W. Kriwacki, J. D. Forman-Kay, Dynamic protein interaction networks and new structural paradigms in signaling. *Chem. Rev.* **116**, 6424-6462 (2016).
21. X. Yin, C. Giap, J. S. Lazo, E. V. Prochownik, Low molecular weight inhibitors of Myc–Max interaction and function. *Oncogene* **22**, 6151-6159 (2003).
22. A. V. Follis, D. I. Hammoudeh, H. Wang, E. V. Prochownik, S. J. Metallo, Structural rationale for the coupled binding and unfolding of the c-Myc oncoprotein by small molecules. *Chem. Biol.* **15**, 1149-1155 (2008).
23. D. M. Clausen *et al.*, In vitro cytotoxicity and in vivo efficacy, pharmacokinetics, and metabolism of 10074-G5, a novel small-molecule inhibitor of c-Myc/Max dimerization. *J. Pharmacol. Exp. Ther.* **335**, 715-727 (2010).
24. M. A. Reed *et al.* (Google Patents, 2018).
25. S. Gil *et al.*, NMR Spectroscopic Studies of Intrinsically Disordered Proteins at Near-Physiological Conditions. *Angew. Chem. Int.* **52**, 11808-11812 (2013).
26. M. Bonomi, G. T. Heller, C. Camilloni, M. Vendruscolo, Principles of protein structural ensemble determination. *Curr. Opin. Struct. Biol.* **42**, 106-116 (2017).

27. G. T. Heller *et al.*, Sequence specificity in the entropy-driven binding of a small molecule and a disordered peptide. *J. Mol. Biol.* **429**, 2772-2779 (2017).
28. M. Bonomi, C. Camilloni, M. Vendruscolo, Metadynamic metainference: Enhanced sampling of the metainference ensemble using metadynamics. *Sci. Rep.* **6**, 31232 (2016).
29. M. Bonomi, C. Camilloni, A. Cavalli, M. Vendruscolo, Metainference: a Bayesian inference method for heterogeneous systems. *Sci. Adv.* **2**, e1501177 (2016).
30. M. Friedemann *et al.*, Effect of methionine-35 oxidation on the aggregation of amyloid-β peptide. *Biochem. Biophys. Rep.* **3**, 94-99 (2015).
31. C. Camilloni *et al.*, Towards a structural biology of the hydrophobic effect in protein folding. *Sci. Rep.* **6**, 1-9 (2016).
32. V. M. Burger, D. J. Arenas, C. M. Stultz, A structure-free method for quantifying conformational flexibility in proteins. *Sci. Rep.* **6**, 29040 (2016).
33. S. I. Cohen, M. Vendruscolo, C. M. Dobson, T. P. Knowles, Nucleated polymerization with secondary pathways. II. Determination of self-consistent solutions to growth processes described by non-linear master equations. *J. Chem. Phys.* **135**, 08B611 (2011).
34. T. P. Knowles *et al.*, An analytical solution to the kinetics of breakable filament assembly. *Science* **326**, 1533-1537 (2009).
35. D. M. Walsh *et al.*, A facile method for expression and purification of the Alzheimer's disease-associated amyloid β-peptide. *FEBS J.* **276**, 1266-1281 (2009).
36. P. Arosio, M. Vendruscolo, C. M. Dobson, T. P. Knowles, Chemical kinetics for drug discovery to combat protein aggregation diseases. *Trends Pharmacol. Sci.* **35**, 127-135 (2014).
37. O. J. Rolinski, M. Amaro, D. J. Birch, Early detection of amyloid aggregation using intrinsic fluorescence. *Biosens. Bioelectron.* **25**, 2249-2252 (2010).
38. S. De *et al.*, Different soluble aggregates of Aβ42 can give rise to cellular toxicity through different mechanisms. *Nat. Commun.* **10**, 1541 (2019).
39. F. Ruggeri *et al.*, Nanoscale studies link amyloid maturity with polyglutamine diseases onset. *Sci. Rep.* **6**, 31155 (2016).
40. F. S. Ruggeri *et al.*, Influence of the β-sheet content on the mechanical properties of aggregates during amyloid fibrillization. *Angew. Chem. Int.* **54**, 2462-2466 (2015).
41. F. S. Ruggeri *et al.*, Identification and nanomechanical characterization of the fundamental single-strand protofilaments of amyloid α-synuclein fibrils. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 7230-7235 (2018).

42. G. Meisl *et al.*, Molecular mechanisms of protein aggregation from global fitting of kinetic models. *Nat. Protoc.* **11**, 252 (2016).
43. T. C. Michaels *et al.*, Thermodynamic and kinetic design principles for protein aggregation inhibitors. *bioRxiv*, (2020).
44. B. Mannini *et al.*, Stabilization and characterization of cytotoxic Aβ40 oligomers isolated from an aggregation reaction in the presence of zinc ions. *ACS Chem. Neurosci.*, (2018).
45. R. Kayed *et al.*, Fibril specific, conformation dependent antibodies recognize a generic epitope common to amyloid fibrils and fibrillar oligomers that is absent in prefibrillar oligomers. *Mol. Neurodegener.* **2**, 18 (2007).
46. The PLUMED consortium, Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **16**, 670-673 (2019).
47. E. Lescop, P. Schanda, B. Brutscher, A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. *J. Magn. Reson.* **187**, 163-169 (2007).
48. B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4**, 435-447 (2008).
49. S. Piana, K. Lindorff-Larsen, D. E. Shaw, How robust are protein folding simulations with respect to force field parameterization? *Biophys. J.* **100**, L47-L49 (2011).
50. A. D. MacKerell Jr *et al.*, All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586-3616 (1998).
51. G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, (2007).
52. H. J. Berendsen, J. v. Postma, W. F. van Gunsteren, A. DiNola, J. Haak, Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684-3690 (1984).
53. M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182-7190 (1981).
54. B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463-1472 (1997).
55. U. Essmann *et al.*, A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577-8593 (1995).
56. K. Vanommeslaeghe, A. D. MacKerell Jr, Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *J. Chem. Inf. Model.* **52**, 3144-3154 (2012).

57. C. G. Mayne, J. Saam, K. Schulten, E. Tajkhorshid, J. C. Gumbart, Rapid parameterization of small molecules using the force field toolkit. *J. Comput. Chem.* **34**, 2757-2770 (2013).

58. P. Robustelli, K. Kohlhoff, A. Cavalli, M. Vendruscolo, Using NMR chemical shifts as structural restraints in molecular dynamics simulations of proteins. *Structure* **18**, 923-933 (2010).

59. J. Pfaendtner, M. Bonomi, Efficient sampling of high-dimensional free-energy landscapes with Parallel Bias Metadynamics. *J. Chem. Theory Comput.* **11**, 5062-5067 (2015).

60. A. Barducci, G. Bussi, M. Parrinello, Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* **100**, (2008).

61. P. Raiteri, A. Laio, F. L. Gervasio, C. Micheletti, M. Parrinello, Efficient reconstruction of complex free energy landscapes by multiple walkers metadynamics. *J. Phys. Chem. B* **110**, 3533-3539 (2006).

62. M. Novo, S. Freire, W. Al-Soufi, Critical aggregation concentration for the formation of early Amyloid-β (1–42) oligomers. *Sci. Rep.* **8**, 1783 (2018).

63. X. Daura *et al.*, Peptide folding: when simulation meets experiment. *Angew. Chem. Int.* **38**, 236-240 (1999).

64. T. Löhr, A. Jussupow, C. Camilloni, Metadynamic metainference: Convergence towards force field independent structural ensembles of a disordered peptide. *J. Chem. Phys.* **146**, 165102 (2017).

65. D. Branduardi, G. Bussi, M. Parrinello, Metadynamics with adaptive Gaussians. *J. Chem. Theory Comput.* **8**, 2247-2254 (2012).

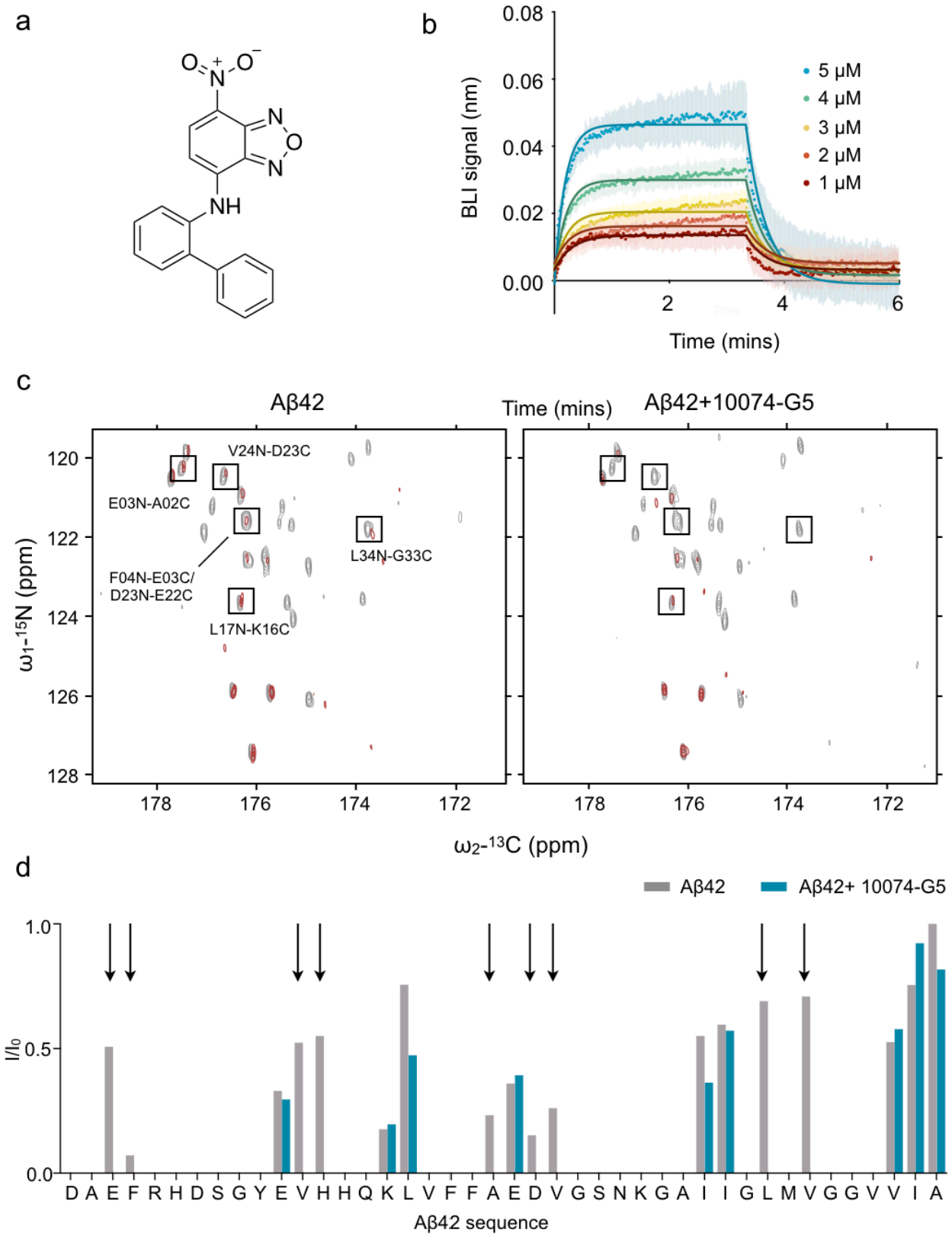
a
b
BLI signal (nm)
Time (mins)
5 μM
4 μM
3 μM
2 μM
1 μM
c
Aβ42
Aβ42+10074-G5
Time (mins)
V24N-D23C
E03N-A02C
L34N-G33C
F04N-E03C/
D23N-E22C
L17N-K16C
ω1-15N (ppm)
ω2-13C (ppm)
d
Aβ42
Aβ42+ 10074-G5
I/I0
D A E F R H D S G Y E V H H Q K L V F F A E D V G S N K G A I I G L M V G G V V I A
Aβ42 sequence

**Figure 1. Characterisation of the interaction of 10074-G5 with monomeric Aβ42**. **(a)** Structure of biphenyl-2-yl-(7-nitro-benzo[1,2,5]oxadiazol-4-yl)-amine, also known as 10074-G5. **(b)** Biolayer interferometry measurements showing the dose-dependent binding of 10074-G5 to an Aβ42-functionalised surface at various concentrations of the added compound. The curves were corrected for baseline drift. Raw data are shown in **Figure S1a**. Control curves showing negligible non-specific binding are shown in **Figure S1b, d.** Global fitting to simple one-phase association and dissociation equations yields association ($k_{on}$) and dissociation ($k_{off}$) rates to be 8.5 x $10^3$ ± 0.2 x $10^3$ $M^{-1}sec^{-1}$ and $k_{off}$ to be 4.7 × $10^{-2}$ ± 2 × $10^{-4}$ $sec^{-1}$, respectively, corresponding to a binding dissociation constant ($K_D$) of 6 μM. For this fit, all five curves were constrained to single, shared $k_{on}$ and $k_{off}$ values. The global $R^2$ for the fits is 0.98. **(c)** 2D $H^{N-BEST}CON$ spectra in the absence (left) and presence (right) of 1:2 Aβ42:10074-G5 with (red) and without (grey) selective water pre-saturation, performed at 15 °C. **(d)** Quantification of the relative $I/I_0$ intensities from (c) shows that the peptide amide groups are more exposed to solvent in the presence of 10074-G5 (blue) as compared to its absence (grey). Arrows highlight regions along the sequence in which signals are detectable in the absence of the compound, but not in its presence, thus suggesting that 10074-G5 increases the solvent exposure of specific regions of Aβ42.

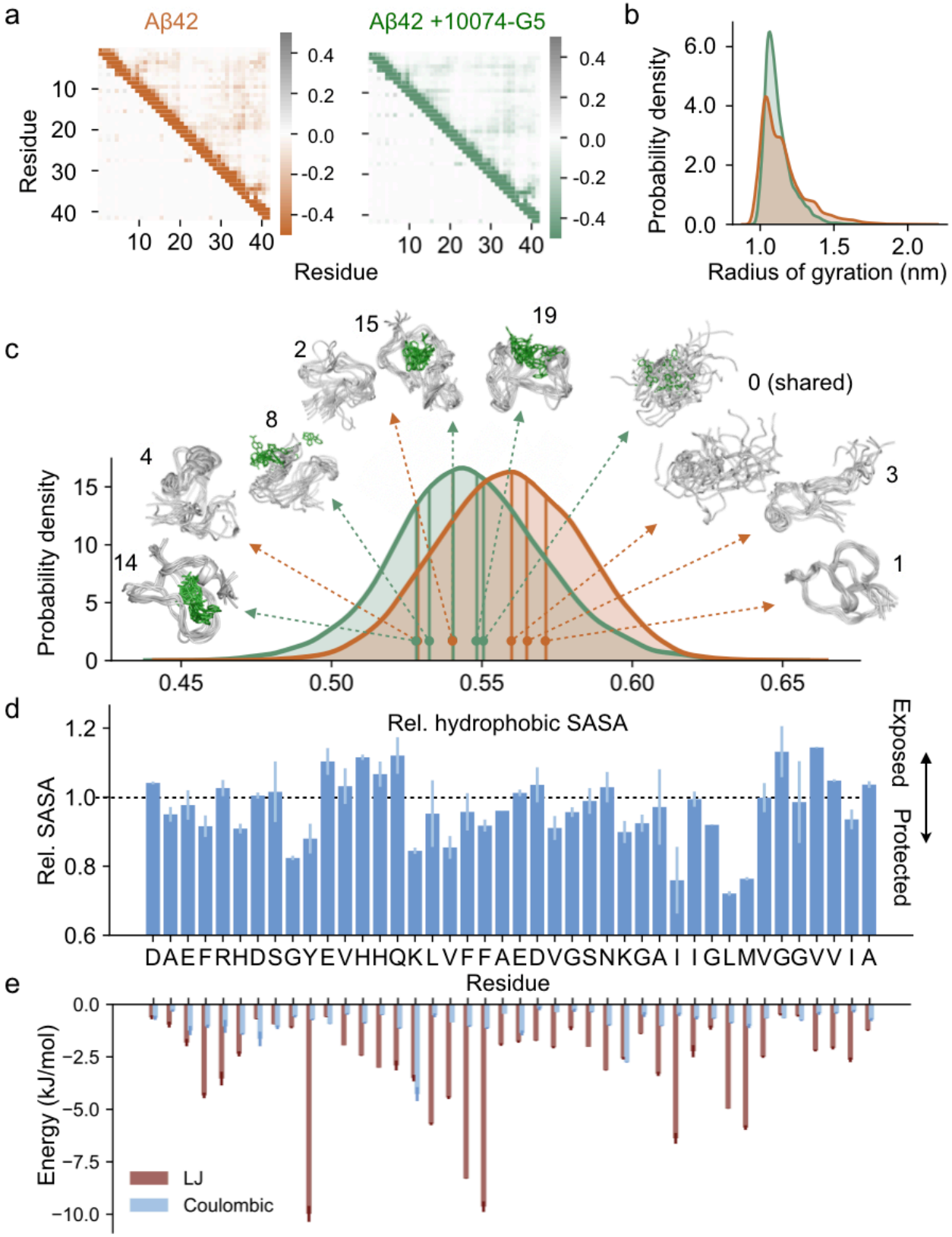
a
Aβ42
Aβ42 +10074-G5
Residue
10
20
30
40
0.4
0.2
0.0
-0.2
-0.4
b
Probability density
6.0
4.0
2.0
0.0
1.0
1.5
2.0
Radius of gyration (nm)
c
15
19
2
0 (shared)
8
4
3
14
1
Probability density
0.45
0.50
0.55
0.60
0.65
d
Rel. hydrophobic SASA
Rel. SASA
1.2
1.0
0.8
0.6
Exposed
Protected
D A E F R H D S G Y E V H H Q K L V F F A E D V G S N K G A I I G L M V G G V V I A
Residue
e
Energy (kJ/mol)
0.0
−2.5
−5.0
−7.5
−10.0
LJ
Coulombic

**Figure 2. Metadynamic metainference simulations characterise the dynamic binding and show how 10074-G5 promotes Aβ42 conformations with less hydrophobicity. (a)** Metadynamics metainference simulations demonstrate that inter-residue contact maps for Lennard-Jones (upper right) and Coulomb (lower left) potentials for the unbound (orange) and the bound (green) structural ensembles of Aβ42 with 10074-G5 are highly similar. **(b)** Radii of gyration for the unbound and bound structural ensembles are also highly similar (shown using kernel density estimates of 35,000 points each sampled based on metadynamics weights using a Gaussian kernel). **(c)** Relative hydrophobic solvent accessible surface area (SASA) of Aβ42 (total hydrophobic area over total surface area) of the bound and unbound ensembles, showing that 10074-G5 decreases the relative exposed hydrophobicity of Aβ42. The holo ensemble was calculated only on the protein surface, but accounts for the presence of the compound. Data are shown using kernel density estimates as described in (b). Some of the representative structures of Aβ42 (orange) and 10074-G5 (green) from within these distributions are shown. Numbers indicate cluster IDs shown in **Figure 3**. **(d)** Ratio (bound/unbound) of the ensemble-averaged, total SASA per residue showing regions of Aβ42 that become more exposed or protected in the presence of 10074-G5. SASAs of the bound ensemble were calculated on the protein, in the presence of 10074-G5. **(e)** Ensemble-averaged, residue-specific Lennard-Jones (LJ) and Coulomb interaction energies show that 10074-G5 has strong interactions with aromatic and charged residues. Error bars represent standard deviations between first and second halves of the analysed trajectories in (d) and (e).

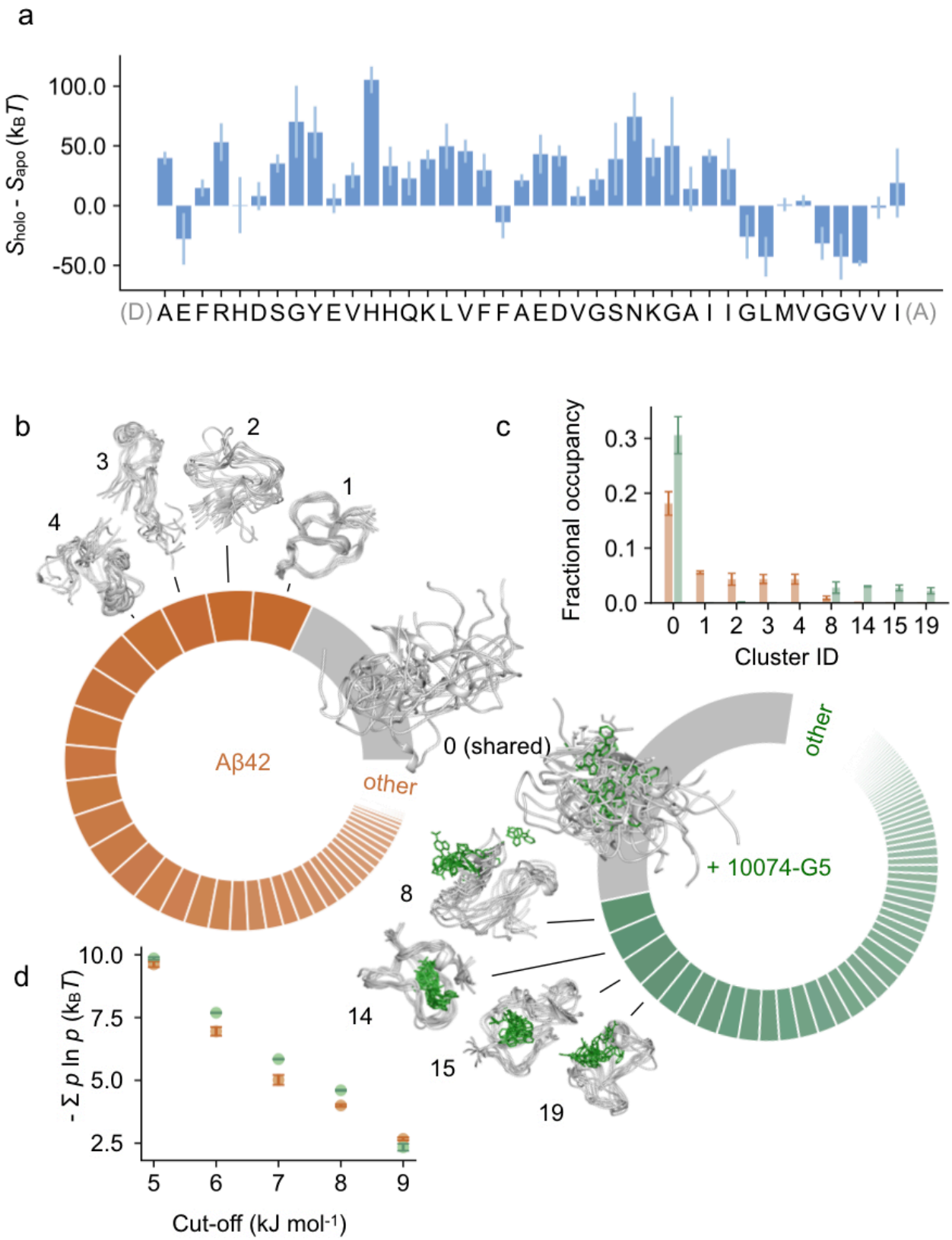
a
$S_{holo}$ - $S_{apo}$ ($k_B T$)
100.0
50.0
0.0
-50.0
(D) A E F R H D S G Y E V H H Q K L V F F A E D V G S N K G A I I G L M V G G V V I (A)
b
1
2
3
4
Aβ42
other
0 (shared)
other
+ 10074-G5
8
14
15
19
c
Fractional occupancy
0.3
0.2
0.1
0.0
0 1 2 3 4 8 14 15 19
Cluster ID
d
- Σ p ln p ($k_B T$)
10.0
7.5
5.0
2.5
5 6 7 8 9
Cut-off (kJ mol-1)

**Figure 3. 10074-G5 increases the conformational entropy of Aβ42. (a)** Residue-specific differences in the conformational entropies, *S*, between the holo and apo ensembles, estimated from normalised, two-dimensional Ramachandran histograms (100 × 100 bins) of each residue using $S = -\sum b \ln b$ where $b$ is the occupancy of a given bin. **(b)** Donut plots quantifying the conformational states of Aβ42 in the unbound (left, orange) and bound (right, green) simulations. Clustering was performed on concatenated trajectories, considering only Aβ42. Inter-residue contact maps based on the Lennard-Jones potential were used as input for GROMOS clustering (*63*). The cut-off value is 8.5 kJ $mol^{-1}$. Each slice represents a distinct state. The simulations share one major state (Cluster 0, grey), which comprises 18 and 31% of the unbound and bound ensembles, respectively. **(c)** Convergence of the 5 most populated clusters. Bar plot shows fractional cluster occupancies for the unbound (orange) and bound (green) simulations. Fractional cluster occupancies were calculated on 35,000 frames for each concatenated trajectory sampled based on metadynamics weights. **(d)** The conformational entropy of Aβ42, estimated via Gibbs entropy, is consistently higher in the 10074-G5-bound form of the peptide for several clustering cut-off values. The conformational entropy was calculated such that the weights, *p*, of each state correspond to the fractional occupancy as determined by the GROMOS clustering algorithm (*63*). Error bars represent ± standard deviations of values calculated from the first and second halves of the simulations in (c) and (d).

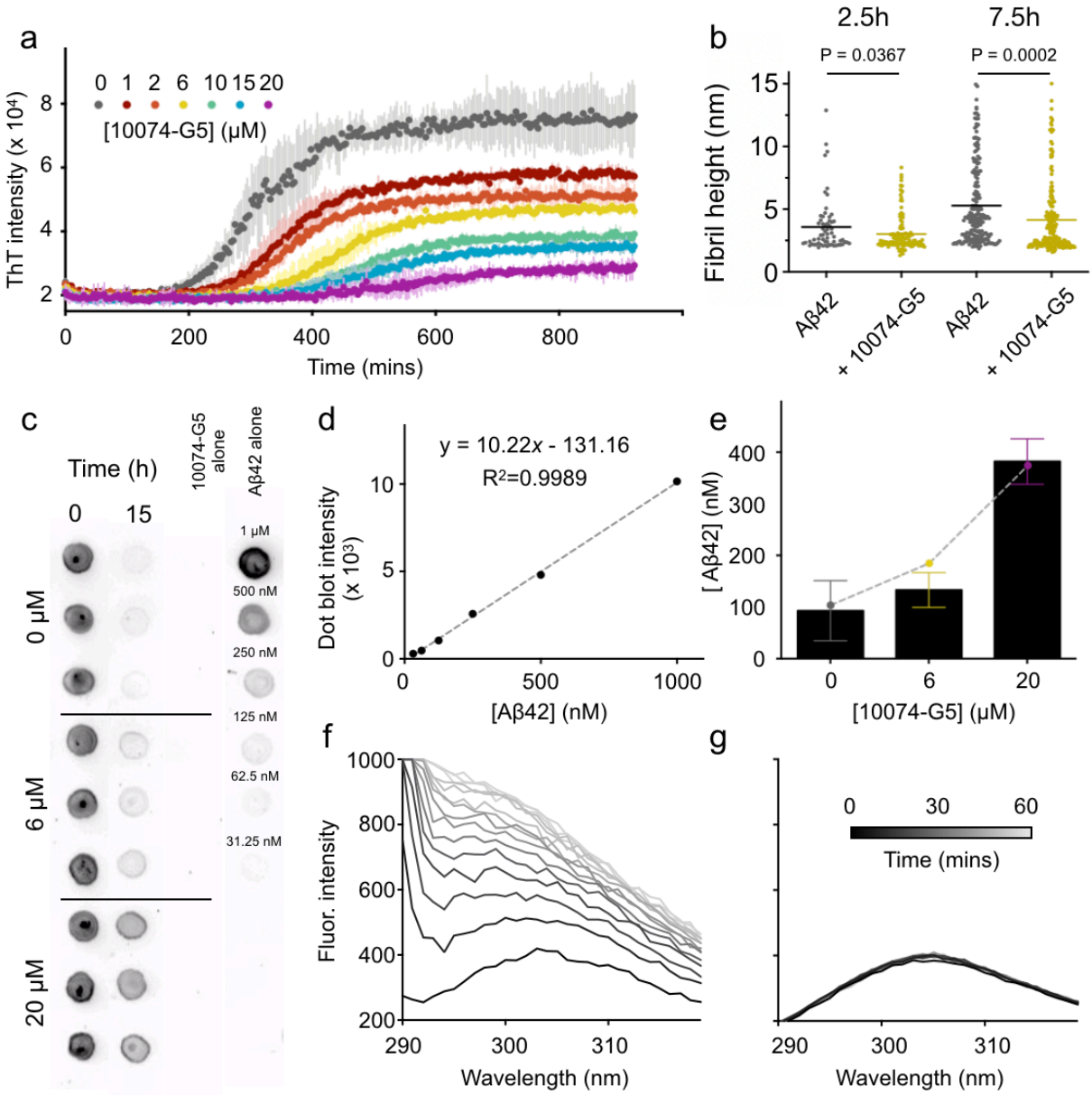


**Figure 4. 10074-G5 sequesters monomeric Aβ42 and inhibits its aggregation**. **(a)** ThT measurements using 1 μM Aβ42 show concentration-dependent effects of 10074-G5 on Aβ42 aggregation. Measurements were taken in triplicate. The concentration of DMSO was held constant across all samples. **(b)** Distributions of cross-sectional heights of 1 μM Aβ42 fibrils at 2.5 h (N ≥ 60) and 7.5 h (N ≥ 200) formed with and without 6 μM 10074-G5, from single-molecule analyses of AFM maps (**Figure S7**). Lines indicate means. P values were determined by unpaired, two-tailed Student's t-test. Fibrillar aggregates formed in the presence of 10074-G5 have smaller cross-sectional diameters than those formed in its absence. **(c)** Dot blot of soluble Aβ42 before and after the aggregation of 1 μM Aβ42 with and without 10074-G5 using the W0-2 antibody indicates sequestration of soluble Aβ42. Blotting was performed in triplicate.

Fit **(d)** and quantification **(e)** used to estimate the concentration of soluble Aβ42 remaining at the end of the aggregation reaction from (c). The dashed line in (e) represents fit to Eq. S13, describing the equilibrium concentration of unreacted monomer from a competitive binding of free monomers to fibril ends and inhibitor. Using this simple fit, we determined the fitted affinity of 10074-G5 for the soluble material to be 7 ± 1 μM. Intrinsic fluorescence profiles of Tyr10 of 5 μM Aβ42 in the absence **(f)** and presence **(g)** of 1:1 10074-G5 over 1 h show that 10074-G5 delays an increase in fluorescence, suggesting 10074-G5 inhibits early aggregation events including oligomerisation and multimerization. Error bars represent ± standard deviations in (a) and (e).

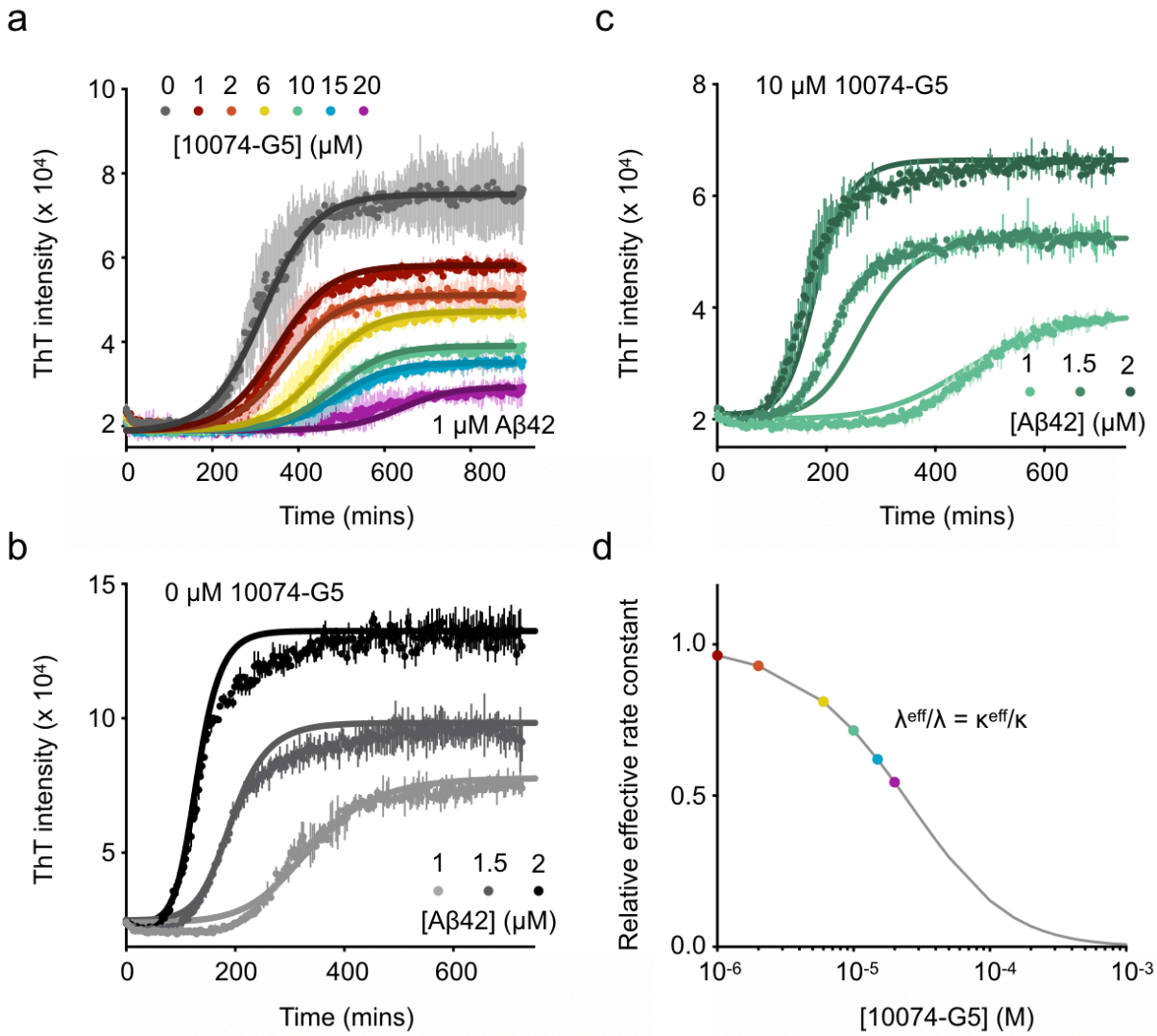


**Figure 5. 10074-G5 inhibits Aβ42 aggregation primarily by monomer sequestration. (a)** Global fit of ThT kinetic curves to a monomer sequestration model (Eq. S11), in which 10074-G5 affects the aggregation by binding free monomers. Measurements were taken in quintuplicate, and the concentration of DMSO was held constant across all concentrations of 10074-G5. The theoretical curves are obtained using Eq. S10 with unperturbed kinetic obtained from (b) leaving $K_D$ as the only global fitting parameter (Eq. S14). The global fit yields $K_D =$ 40 μM. Global fits to Eq. S14 were performed on normalised data to extract changes in the rate parameters in the presence of 10074-G5. (**b**) Global fit to Eq. S10 of ThT kinetic traces of the aggregation reaction for increasing concentrations of Aβ42 (1, 1.5 and 2 μM) in the absence of 10074-G5. Measurements were taken in duplicate or triplicate. (**c**) Overlay of theoretical kinetic curves from (a) with independent ThT kinetic traces of the aggregation reaction for increasing concentrations of Aβ42 (1, 1.5 and 2 μM) in the presence of 10 μM 10074-G5. Solid curves are

predictions of the kinetic monomer sequestration model using the same rate parameters and inhibitor binding constant as in (a) and no fitting parameters. Measurements were taken in duplicate or triplicate. **(d)** Effective rates of aggregate proliferation through primary (λ) and secondary (κ) nucleation in the presence of varying concentrations of 10074-G5 determined using the global fit in (a). Error bars represent ± standard deviations in (a-c).

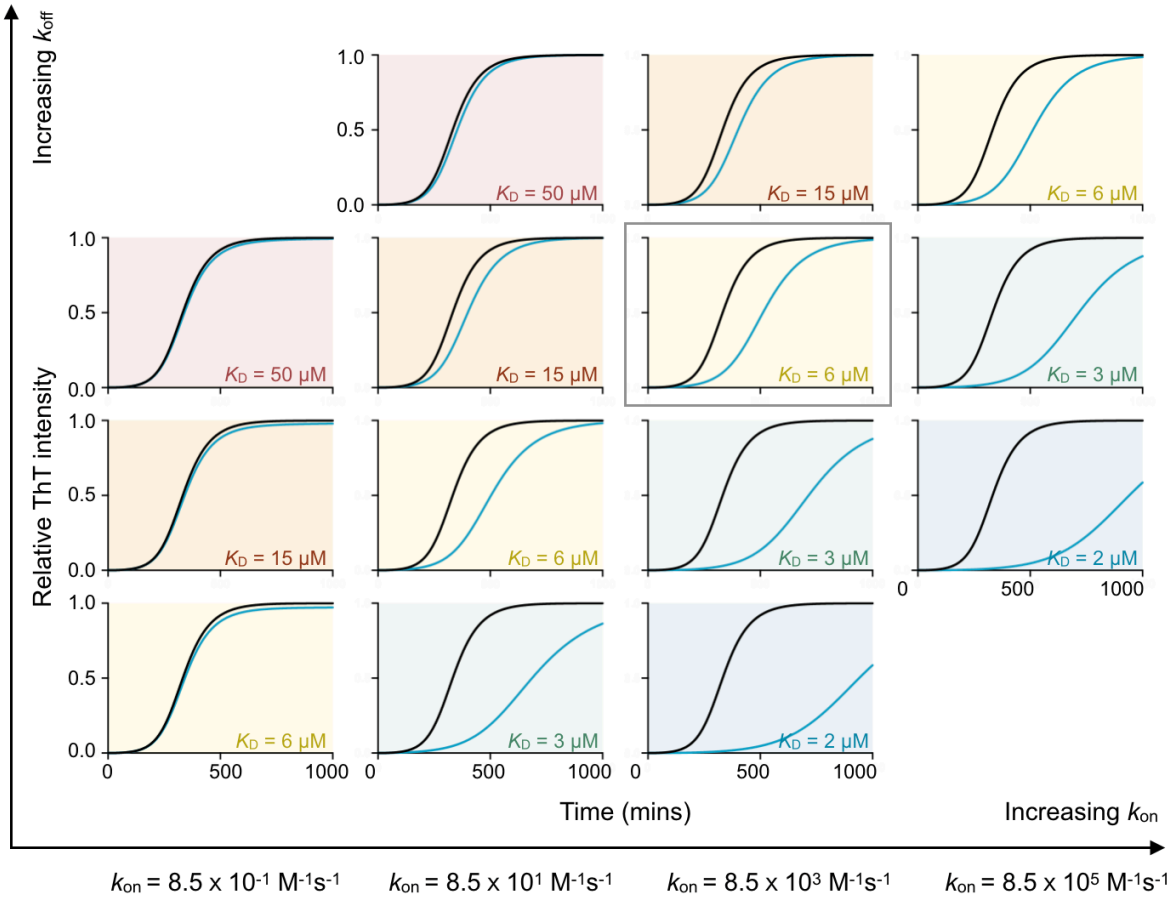


**Figure 6. Phase diagram illustrating numerical solutions to the kinetic equations for different $k_{on}$ and $k_{off}$ rates.** Curves represent kinetic aggregation traces of 1 μM Aβ42 in the absence (black) and presence of 2 μM of compound (blue). Diagonals correspond to constant values of $K_D$. The values of $k_+k_n$ and $k_2k_+$ are the same as those shown in **Figure 5a**.