Mechanisms of small-molecule binding to intrinsically disordered proteins

Citation for published version:
Cuchillo, R & Michel, J 2012, 'Mechanisms of small-molecule binding to intrinsically disordered proteins' Biochemical Society Transactions, vol. 40, no. 5, pp. 1004-1008. DOI: 10.1042/BST20120086

Digital Object Identifier (DOI):
10.1042/BST20120086

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Biochemical Society Transactions

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Mechanisms of small molecule binding to intrinsically disordered proteins

Rémi Cuchillo and Julien Michel

EaStCHEM School of Chemistry, University of Edinburgh, Edinburgh, EH9 3JJ, United Kingdom

Received March 05, 2012

Abstract
Intrinsically Disordered Proteins (IDPs) play crucial roles in many important cellular processes such as signalling or regulation and are attractive therapeutic targets for several diseases. The considerable structural flexibility of IDPs poses a challenge for rational drug discovery approaches. Consequently structure-based drug design efforts to date have mostly focused on inhibiting interactions of IDPs with other proteins whose structure can be solved by conventional biophysical methods. Yet, in recent years several examples of small molecules that bind to monomeric IDPs in their disordered states have been reported, suggesting that this approach may offer new opportunities for therapeutic interventions. Further developments of this strategy will greatly benefit from an improved understanding of molecular recognition mechanisms between small molecules and IDPs. This article summarizes findings from experimental and computational studies of the mechanisms of interaction between small molecules and three IDPs in their disordered states: c-Myc, Aβ peptide and α-synuclein.

* To whom correspondence should be addressed (email julien.michel@ed.ac.uk)

Abbreviations used: IDPs, intrinsically disordered proteins; Aβ, Alzheimer β-amyloid peptide; bHLHZip, basic helix-loop-helix leucine zipper; AD, Alzheimer’s disease; APP, Amyloid precursor protein; DA, dopamine; MD, molecular dynamics.

Keywords: intrinsically disordered proteins, small molecules, c-Myc, α-synuclein, Aβ peptide.
Introduction

The cell machinery is controlled by a large number of interactions between proteins and nucleic acids. It is now well appreciated that a large number of proteins do not adopt a single well defined structure under native conditions. Proteins that contain a segment of at least 30 consecutive disordered residues in their native state are typically classified as Intrinsically Disordered Proteins (IDPs) [1]. In comparison with globular proteins, IDPs tend to contain fewer hydrophobic residues but are generally enriched in charged amino acids [2-4]. IDPs can adopt a broad range of conformations, ranging from collapsed to fully extended. The considerable flexibility of IDPs facilitates interactions with a broad range of proteins and explains why IDPs often play key roles in important cellular processes such as signalling or transcription [5,6]. Molecular recognition between an IDP and a partner protein can involve a disorder-to-order transition through a coupled folding upon binding mechanism, which produces high-specificity low-affinity complexes [7]. There are however several examples of IDPs that remain disordered upon complex formation [8].

IDPs are attractive therapeutic targets as they are often implicated in a broad range of diseases such as cancers, cardiovascular disease or neurodegenerative diseases. However the considerable flexibility of IDPs presents a challenge for drug discovery approaches [9]. Due to their lack of a well-defined tertiary structure, it is generally not possible to determine the structure of isolated IDPs. So far structure-based approaches to inhibiting IDPs have targeted either partner proteins that are ordered, or ordered complexes, in those cases where IDPs fold upon binding. For instance, the p53 tumor suppressor is an IDP that is involved in the progression of more than 50% of human cancers. The transcriptional activity of p53 is tightly regulated by partner protein MDM2 and cancer cells often over express MDM2 to inhibit p53 function [10]. As the p53 binding domain of MDM2 is folded, crystal structures can be readily obtained and have been exploited to design several classes of small molecule inhibitors of p53/MDM2 [11]. Some of the most successful inhibitors have advanced in clinical trials [12].

However several protein-protein interactions involve two IDPs whose structure cannot be solved in isolation. Even in those instances where two IDPs mutually fold upon binding, the structure of the complex may not reveal pockets that small molecules could readily bind to. Thus a more general route to inhibiting IDP function would be to directly target their disordered state with small molecules. Historically, this approach has not been considered feasible [13]. However this view has been challenged in recent years, with the discovery of several small molecules that inhibit IDP function by binding to their unfolded state [14-16]. The interactions of small molecules with IDPs challenge our understanding of molecular recognition and it is important to clarify the mechanisms of IDP-small molecule interactions before such proteins can be more routinely targeted. This review article focuses on three well-studied systems, the oncoprotein c-Myc, the Alzheimer β-amyloid peptide (Aβ) and α-synuclein.

c-Myc

The proto-oncogene protein c-Myc is constituted of 439 amino acids and contains an 88 amino acids basic helix-loop-helix leucine zipper (bHLHZip) domain. In its monomeric form, c-Myc is intrinsically disordered. c-Myc has been shown to interact with a large number of other proteins. The specific interaction between c-Myc and the protein Max has been studied extensively because the c-Myc/Max heterodimer binds DNA and regulate gene expressions [17]. It has been shown that overexpression of c-Myc is frequent in many cancers, and disruption of the c-Myc/Max interaction is a possible anticancer strategy [5].

Structurally diverse small molecules inhibiting the formation of this complex were discovered through a yeast-two hybrid screen [14]. Biophysical studies using fluorescence assays, NMR and circular dichroism measurements were performed to characterize protein-ligand interactions [18-20]. These
studies suggest that the small molecules disrupt the c-Myc/Max interaction by stabilizing conformations in monomeric c-Myc that are incompatible with heterodimerization with Max. Three distinct binding sites, encompassing residues 366-375, 375-385 and 402-409, have been mapped on the c-Myc bHLHZip domain [18]. Remarkably, the three distinct c-Myc binding sites can be occupied simultaneously by different ligands. These results suggest that the c-Myc/small molecules interactions are fairly localized and can be predicted from primary sequence analysis. Indeed, protein disorder prediction algorithms can locate approximately the c-Myc small molecule binding sites, which tend to be enriched in hydrophobic amino-acids in comparison with the rest of the domain [20]. In addition, many of the small molecule ligands can bind truncated c-Myc segments containing a single binding site with a binding affinity similar to the full c-Myc bHLHPZip domain. For instance, the small molecule 10058-F4 binds in a fluorescence polarisation assay c-Myc353-437 with a $K_d$ of 5.3 ± 0.7 $\mu$M and c-Myc402-412 with a $K_d$ of 13.3 ± 1 $\mu$M [20]. Furthermore, similar chemical shift perturbations were observed for c-Myc353-437 and c-Myc402-412 upon binding 10058-F4. NMR and circular dichroism studies suggest that c-Myc remains disordered upon binding 10058-F4. Ligand binding appears to lead to formation of a hydrophobic cluster between the ligand and the side-chains of Tyr$^{402}$, Ile$^{403}$, Leu$^{404}$ and Val$^{406}$ (Figure 1). Molecular dynamics studies performed in our group reveals multiple distinct binding modes for 10058-F4, with frequent stacking interactions with Tyr$^{402}$ as well as hydrogen-bonding interactions of with the main chain of Tyr$^{402}$, Val$^{406}$ and Lys$^{412}$ (Unpublished work, J Michel and R Cuchillo).

**Alzheimer β-amylloid peptide**

Alzheimer disease (AD) is a neurodegenerative pathology characterized by the formation of senile plaques into the brain [21]. The aggregation of Aβ, also called Amyloid β peptide, is known to be one of the main components of those plaques and may be associated with the pathogenesis of AD [22,23]. Aβ (36-43 amino acids) is produced by the successive cleavage of the amyloid precursor protein (APP) by the enzymes β-secretase and γ-secretase. Although the role of APP is not completely characterized, it appears to be crucial for synapse formation and function [24]. The aggregation of this small peptide, as well as with other compounds such as apolipoprotein E, induces the development of senile plaques. The Aβ peptide adopts a folded helical structure in membrane environments, but an aggregation prone β-sheet conformation in aqueous solution [25].

During the last decades, many peptide and small molecule inhibitors of Aβ aggregation have been discovered, primarily through in vitro assays [26]. Current small molecules inhibitors appear to inhibit Aβ aggregation through at least two distinct mechanisms. For instance scylo-inositol derivatives have been shown by electron microscopy experiments to bind and stabilize monomeric and trimeric forms, thus blocking aggregation [27,28]. On the other hand compounds like Thioflavin T or Congo red appear to interact with Aβ peptide aggregates, although decades of studies on these compounds have produced several conflicting models describing binding mechanisms. Groening has recently reviewed extensively plausible hypotheses [29].

Computational studies have attempted to clarify protein-ligand interactions. Molecular dynamics simulations were recently performed for ten small molecule inhibitors in presence of a truncated Aβ (Aβ12-28) [30]. Although the small molecules did not exhibit a predominant binding modes and did not dramatically affect the secondary structure preferences of Aβ12-28, a number of conserved interactions with Aβ12-28 could be observed. Most of the ligands interacted preferentially with the N-terminal portion of the peptide (residues 13-20). Energetic analysis revealed favorable electrostatic interactions with several charged amino acids (His$^{13}$, His$^{14}$, Lys$^{16}$). Additionally, favorable hydrophobic interactions are observed between the inhibitors and the entire N-terminals stretch, the sites of highest interaction probability are near the side chains of Phe$^{19}$ and Phe$^{20}$. The binding affinities appear to be roughly correlated with the number of aromatic groups and charged groups present in the ligands. MD simulations have also been performed to examine the interactions of two small molecules, Pep1b and Dec-DETA that were designed to stabilize the central helix in the Aβ peptide [31]. Both ligands appear
to stabilize the Aβ central helix (residues 15-24) in Aβ_{13-26} by interacting preferentially with two charged amino acids, Glu^{22} and Asp^{23}. In addition, electrostatic interactions with His^{13}, Lys^{16} as well as hydrophobic interactions with Phe^{19} and Phe^{20} were also reported for Pep1b (Figure 1). It appears that the extended side-chains interactions between the small molecules and the Aβ peptide disfavor intramolecular side-chain interactions that would destabilize the central α-helix. Recently, molecular dynamics simulations were used to study the interactions of inositol ligands with (Gly-Ala)$_4$ modelled either as small disordered or b-sheet aggregates of 4 peptides, or an extended fibril like oligomer. The ligands were observed to form predominantly mono-dentate and bi-dentate hydrogen bonds with the peptide backbone. The results suggested that inositol does not inhibit amyloid formation by dispersing preformed aggregates, but more likely binds to the surface of prefibrillar aggregates. The computed dissociation constants of the ligands were much higher than the observed in vitro inhibitory concentrations of Aβ peptide aggregation, suggesting the existence of important side-chain interactions with Aβ peptide aggregates. [32].

α-synuclein

The 140 amino acids protein α-synuclein is constituted of three distinct domains. The central region of α-synuclein is known to be crucial for the aggregation of α-synuclein fibrils, one of the main component of Lewy bodies associated with many neurodegenerative diseases such as Parkinson’s Disease (PD) [33,34]. Under physiological conditions, α-synuclein normally adopts an helical conformation that is non-pathogenic and plays a role in neurotransmitter release. It is still not well understood how α-synuclein first forms soluble oligomers called protofibrils, followed by the development of β-sheet rich α-synuclein fibrils. In light of these observations, a deeper molecular-level understanding of interactions between monomeric, protofibril and fibril forms is important to facilitate the discovery of small molecules inhibitors of α-synuclein fibrillization.

A few years ago, fifteen fibrillization inhibitors were found by screening a small molecule library using a fibrillization assay [35]. Many of these inhibitors are members of the catecholamine family and include Dopamine (DA). There is controversy about the mechanisms of interactions between DA and α-synuclein. Conway et al. have suggested that DA readily oxidizes into dopamine-derive orthoquinone (DAQ) that subsequently form a covalent adduct with α-synuclein by radical coupling to form dityrosine linkages or nucleophilic attack of a lysine side chain [35]. On the other hand Norris et al. failed to detect significant levels of DA-syn-α adducts and suggested instead that binding occurs through non covalent interactions with the α-synuclein segment 125Tyr-Glu-Met-Pro-Ser^{129} [36]. Herrera et al. used docking calculations and molecular dynamics simulations to study the interactions of dopamine and several plausible oxidised derivatives with an NMR derived structural ensemble of α-synuclein. In the majority of the simulated complexes, the ligands interacted through a broad range of hydrogen bonding and hydrophobic interactions with the region 125Tyr-Glu-Met-Pro-Ser^{129}. Additionally, large electrostatic interactions were computed between the ligands and residue Glu^{83} located in the non-β-amyloid region of α-synuclein. Point mutations to Ala residues in the 125YEMPS^{129} region did not prevent DA inhibition of α-synuclein aggregation, suggesting that DA interacts non specifically with this region. On the other hand, mutation of Glu^{83} to Ala strongly impaired the ability of DA to inhibit a-synuclein aggregation. [37].

Non catecholamine inhibitors of α-synuclein have also been identified. A broad range of biophysical methods were used by Lendel et al. to characterize the interactions of Congo red and Lacmoid with α-synuclein. They concluded that these two small molecules interact broadly with the N-terminal and central region of α-synuclein as small oligomeric species [38].

Discussion
Although small molecules have now been found to interact directly with several IDPs in their monomeric form, an important challenge is to clarify the specificity of the interactions. For instance, there are numerous proteins that contain a bHLHZip domain similar to c-Myc. Consequently, several small molecules that inhibit the c-Myc-Max complex also inhibit related bHLHZip pairs. To illustrate, the compound 10058-F4 has also been shown in a yeast two hybrid assay to disrupt the complexes MyoD-E2-2, Mad1-Max and Mxl1-Max, although several other bHLHZip pairs were not inhibited [14]. Several of the Dopamine derivates that inhibit α-synuclein aggregation have also been shown to also dissolve fibrils of the Aβ peptide in vitro [39]. Congo red and Lactoin bind readily to β-synuclein, a protein closely related to α-synuclein which does not aggregate under physiological conditions [38].

In several cases, relatively structurally diverse small molecules have been found to interact with similar regions in an IDP. Additionally, many studies suggest that the complexes between small molecules and IDPs remain disordered [40]. This suggests that the binding of the small molecules is driven by a large number of weak interactions [13]. Arguably, unlike proteins, small molecules are unlikely to induce IDP folding upon binding, as the relatively limited intermolecular contacts they form are unlikely to overcome the large conformational entropy loss necessary to structure an IDP. Structure-based approaches to design ligands for IDPs will therefore have to explicitly consider multiple binding modes.

Although the mechanisms of IDP aggregation are still not well understood, a number of small molecule inhibitors of IDP aggregation have reached clinical studies. For instance, methylthionium chloride, initially developed as an antimalarial agent, has been shown to inhibit in vitro the aggregation of the IDP tau [41]. Results of a phase II clinical trial reported that methylthionium chloride slows down cognitive impairment in patient suffering from AD, thus inhibiting the formation of tau aggregates is a promising strategy for the development of AD treatments [42].

Funding

J.M. is supported by a Royal Society University Research Fellowship. This research was also part funded by a Marie Curie International re-integration grant (PIRG-GA-2010-276785).

Figure legend

**Figure 1.** Summary of the main interactions observed in three IDP/ligand complexes, c-Myc/10058-F4, α-synuclein/Dopamine and Amyloid β peptide/Pep1b.
References


formation and cytotoxicity by hydroxyindole derivatives. Biochemistry. 45, 4727-4735


