Binding of sulphonated indigo derivatives to RepA-WH1 inhibits DNA-induced protein amyloidogenesis

Fátima Gasset-Rosa¹, María Jesús Maté², Cristina Dávila-Fajardo¹, Jerónimo Bravo³ and Rafael Giraldo¹,*

¹Department of Molecular Microbiology, ²Department of Protein Science, Centro de Investigaciones Biológicas (CSIC), C/ Ramiro de Maeztu, 9. E-28040 and ³Structural Biology and Biocomputing Programme, Centro Nacional de Investigaciones Oncológicas (CNIO), C/ Melchor Fernández Almagro, 3. E-28029 Madrid, Spain

Received November 19, 2007; Revised January 9, 2008; Accepted February 1, 2008

ABSTRACT

The quest for inducers and inhibitors of protein amyloidogenesis is of utmost interest, since they are key tools to understand the molecular bases of proteinopathies such as Alzheimer, Parkinson, Huntington and Creutzfeldt–Jakob diseases. It is also expected that such molecules could lead to valid therapeutic agents. In common with the mammalian prion protein (PrP), the N-terminal Winged-Helix (WH1) domain of the pPS10 plasmid replication protein (RepA) assembles in vitro into a variety of amyloid nanostructures upon binding to different specific dsDNA sequences. Here we show that di- (S2) and tetra-sulphonated (S4) derivatives of indigo stain dock at the DNA recognition interface in the RepA-WH1 dimer. They compete binding of RepA to its natural target dsDNA repeats, found at the repA operator and at the origin of replication of the plasmid. Calorimetry points to the existence of a major site, with micromolar affinity, for S4-indigo in RepA-WH1 dimers. As revealed by electron microscopy, in the presence of inducer dsDNA, both S2/S4 stains inhibit the assembly of RepA-WH1 into fibres. These results validate the concept that DNA can promote protein assembly into amyloids and reveal that the binding sites of effector molecules can be targeted to inhibit amyloidogenesis.

INTRODUCTION

Aggregation of proteins into amyloid assemblies is the conundrum of an increasing number of proteinopathies with devastating impact on human health. Although amyloids share related 3D structures, the factors responsible for inducing protein amyloidogenesis are diverse (1). Among the latter, nucleic acid ligands have been extensively studied for the mammalian prion protein (PrP), the causative agent of lethal transmissible spongiform encephalopathies (2). The conformation of the cellular form of this protein (PrPc) is transformed into its pathogenic variant (PrPsc) upon binding to long, mixed DNA or RNA sequences in vitro, promoting its assembly into amyloid fibres and spheroids (3–7). More recently, the interaction of PrP with a short, specific dsDNA sequence has been characterized by NMR spectroscopy (8) and the in vitro selection of small ssDNA thioaptamers binding to PrP reported (9). The latter study points to the existence in PrP of two DNA-binding sites, one with high affinity (specific) and another with lower affinity (non-specific) (9). The reports by Supattapone and co-workers (10), on the strict requirement of polyanionic additives for the successful amplification in vitro of PrPsc from recombinant PrP and on the existence of a stable complex constituted by several PrP molecules bound to RNA/ssDNA (11), have the highest relevance for understanding prion amyloidogenesis.

RepA is a multifunctional DNA binding protein encoded by the Pseudomonas pPS10 plasmid (12). RepA dimers bind to repA gene operator repressing its own transcription, whereas RepA monomers activate plasmid replication after cooperative binding to four directly repeated DNA sequences (termed iterons) (13). Single iteron binding, by itself, enhances dissociation of RepA dimers and induces a structural transformation (14,15) in the N-terminal Winged-Helix dimerization domain (WH1) (16) that implies the conversion of part of its α-helical elements into β-strands and loops (17). Although stable binding of RepA to dsDNA requires the presence of a...
second, C-terminal WH domain (WH2), both WH1 and WH2, when isolated, can bind to their targets at the operator and iteron sequences (16). The conformational changes experienced by RepA, paralleling those in known amyloid forming proteins (1), have recently inspired a search for conditions leading to RepA-WH1 amyloidosis in vitro (18). Specific operator or iteron core sequences (11 bp), combined with point mutations in an amyloidogenic sequence located opposite to the DNA-binding interface (17), drive the assembly of the domain into a range of amyloid nano-structures spanning from irregular aggregates to well ordered fibres, through regular spheroids (18). The key determinant for the final type of assembly or ‘strain’ is the sequence of the dsDNA ligand (operator for the fibres), albeit DNA itself is not a component of the matured amyloids (18). This suggests that transient binding of a specific dsDNA to WH1 would shield an electropositive patch at the DNA binding interface, thus assisting the ordered assembly of the protein into fibres. The latter is achieved through a cross-β spine made of the amyloidogenic peptide sequence L26VLCAVSLI34, distal to the dsDNA-binding site (18).

RepA-WH1 provides a suitable and novel bacterial model system to study sequence-specific DNA-promoted protein amyloidogenesis. Here, we have searched for small molecules docking in silico to RepA-WH1 and found di- (S2) and tetra-sulphonated (S4) derivatives of the classic indigo stain that compete with RepA binding to dsDNA. The thermodynamic characterization of S4-indigo interaction shows that it involves a major binding site in WH1 ($K_d = 0.23 \mu M$). Inhibition by the sulphonated indigoids of RepA-WH1 binding to the operator dsDNA sequence abolishes fibre assembly, constituting a proof of concept for DNA-induced protein amyloidogenesis (18). Unlike the wealth of small molecules described so far to inhibit protein aggregation into amyloids (19,20), those reported here are the first acting on the binding site for a regulator or enhancer of amyloidogenesis, rather than at the peptide sequences directly assembling into the landmark amyloid cross-β structure.

MATERIALS AND METHODS

Docking of sulphonated indigoids to RepA-WH1

The chemical structures of indigo (3,3’-dioxo-2,2’-bisindole) and its 5,5’-disulphonic (indigotine or indigo carmine) and 5,5’,7,7’-tetasulphonic derivatives were outlined by using ChemDraw (v.10.0, CambridgeSoft) and then exported to SMILES code. The 3D structures of the molecules were generated by means of Corina (http://www.molecular-networks.com/online_demos/corina demo.html) and saved as PDB files. Docking of the indigoid molecules to the structure of RepA-WH1 dimer (17; PDB code 1HKQ) was performed using AutoDock (v.3.05) (21), run in an Apple workstation (2 x 2 GHz PowerPC G5, OSX 10.3.9). Being the indigo (bis-indolic) core planar considerably simplifies the flexible docking of the ligands, because it reduces the number of torsional degrees of freedom that have to be taken into account to those concerning the four (S2-indigo) or eight (S4-indigo) substituent sulphate groups. Protein structure was screened within a 50 x 50 x 50 Å cube centred at the dimer interface and sampled with a grid in which points were spaced 0.425 Å. Translation and rotation steps were set to 1.5 Å and 25°, respectively. A Lamarckian genetic algorithm (mutation rate 0.02, crossover rate 0.8) was then used for 25 runs in a population of 50 individuals. All the solutions fitted to the DNA binding interface in RepA-WH1 (17). Independent runs placing the sampled volume around other protein areas did not yield a solution with a noticeable low $\Delta G$.

Protein purification and DNA synthesis

RepA-WT protein and its isolated WH1 domain, either WT or carrying the amyloidogenic enhancing mutation A31V, were purified as described (16,18). Concentrated protein stocks were dialysed against 0.1 M Na$_2$SO$_4$, 20 mM Tris–SO$_4$ pH 8, 4 mM MgSO$_4$, 1 mM 2-MeEtOH, 10% glycerol before long-term storage at −70°C. Protein concentration, determined by A$_{280}$nm (14,16), is expressed in a per-monomer basis. The dsDNA oligonucleotide opsp (5’-CATTGACTTGT/3’-GTAACT GAACA) (18) was synthesized and annealed (0.2 mM stock) as previously described (16). For the DNA binding competition assays by indigoids, PCR was performed using as templates pUC18 derivatives (0.1 mM) carrying either a 39 bp sequence spanning the repA operator or a 22 bp single pPS10-oriV iteron, cloned into the vector Smal site (14,15). The universal primers f17 and r19 (50 pmol each) were used in 40 cycles of amplification by gel-adsorbed Taq DNApol (BioTools). Amplified dsDNAs were purified by means of phenol–chloroform–isoamylic alcohol extraction plus ethanol precipitation.

Electrophoretic mobility shift assays (EMSA)

Ten microlitre aliquots containing 35 pmol of RepA-WT in protein buffer (see above) were incubated for 15 min on ice with serial dilutions of S2/S4-indigo (purchased from TCI Europe, Belgium; 10 mM stocks in water). Then the operator or iteron dsDNAs (calibrated to be ≈50% bound by RepA) were added and incubated for further 15 min on ice, before adding 2.5 μl of 0.05% bromophenol blue in 20% glycerol. Samples were immediately loaded onto 6% polyacrylamide (29:1)-0.25 x TBE minigels. Electrophoresis was run at 150 V, for 1 h at 4°C. After ethidium bromide staining, bands were quantified by UV detection (Bio-Rad Gel Doc 2000).

Isothermal titration calorimetry (ITC)

Six microlitre aliquots of a RepA-WH1-WT stock (150 μM) were extensively dialysed (51 total volume, various changes for 48 h) against 20 mM Tris–HCl pH 8 at 4°C. A quantity of 1.8 ml of the protein solution (a concentration range spanning 75–100 μM was tested in several independent assays) were displayed into a MicroCal VP-ITC (22), degassed for 5–10 min and equilibrated at 16°C. The reference cell contained buffer from the last dialysis passage. Twenty-five injections (10 μl each) of distinct concentrations (0.62–1.41 mM, diluted in dialysis buffer from a 10 mM stock) of S2/S4-indigo were...
performed using the control software of the calorimeter, leaving 150 s between injections to allow for equilibration. Baselines were acquired (the buffer at the sample cell titrated with S2/S4-indigo) and subtracted before integration of data. Calculation of the affinity constants ($K_a$), enthalpy ($\Delta H$) and stoichiometry ($n$) were performed, according to pre-established binding models, by means of non-linear regression best fitting (23; equation 6b therein), as implemented in the Origin-7 software (Microcal).

Transmission electron microscopy (TEM)

RepA-WH1-A31V (35 µM) fibres were assembled in the presence of equimolar amounts of opsp-dsDNA as described (18), with the only modification of including 1-, 3- or 9-fold molar excess of S2/S4-indigo over protein molecules in the incubations (4°C for 1 month). TEM was performed after negative staining (2% uranyl acetate) using a JEOL JEM-1230 electron microscope, operating at 100 kV (18).

RESULTS

Small molecules targeting the dsDNA binding surface in RepA-WH1

The rationale for the selection of molecules inhibiting RepA-WH1 interaction with dsDNA and thus protein amyloidogenesis (18) was the following. Contacts of WH1 with the phosphate backbone of dsDNA are required for the stable binding of RepA to the operator and iteron sequences (16,24). Thus, the electropositive patch defined in the 3D structure of WH1 dimers (17) by the side chains of arginine residues 78, 81 and 91 would account for the $\Delta G_{bind}$ predicted for the unmodified indigo molecule (∼5.54 kcal/mol, data not shown). Models were rendered using PyMOL (http://www.pymol.org).

Figure 1. Sulphonated indigo derivatives target the dsDNA binding interface in RepA-WH1. Left: the most symmetrically arranged docking solutions for the acidic indigo derivatives (cyan) S2 (A) and S4 (B) on the pseudo-2-fold axis of the protein dimer (17; PDB entry 1HKQ; subunits in yellow and green). The calculated values of $\Delta G$ for docking and binding are displayed. Right: close-up views of the sulphonated indigos in electrostatic interaction with the side chains of the Arg residues 78, 81 and 91. The guanidinium group of each Arg residue could make a cation-π interaction with its nearest indol moiety that, together with H-bonding between the side chain of Asn77 and the pyrrolidin keto and amino groups, would account for the $\Delta G_{dock}$ predicted for the unmodified indigo molecule (∼5.54 kcal/mol, data not shown). Models were rendered using PyMOL (http://www.pymol.org).
the pseudo-2-fold symmetry axis in the WH1 dimer (17), the candidate ligand should also exhibit that symmetry. Therefore, a search for small, symmetric, acidic molecules was carried out by visual inspection of the available commercial catalogues. Up to 10 molecules following these criteria, most of them di-carboxylates derived from tartaric acid but also including two sulphonated variants of the blue stain indigo, were tested in silico by docking to the crystal structure of RepA-WH1 dimer (see Materials and Methods section). Those with the best (lowest) ΔG values predicted were the S2 and S4-indigoids (Figure 1), in particular the latter (Figure 1B). The enthalpic term, provided by salt bridges between the electronegative oxygens in the ligand sulphates and the electropositive guanidinium groups in the protein Arg residues, is expected to drive binding, thus tuning the specificity of the interaction. The rigidity of the ligand core leaves the overall entropic term of interaction to rely on the water molecules/ions displaced upon binding.

The S2/S4-indigoids were then purchased and their abilities to inhibit the binding of full-length RepA-WT dimers to operator (Figure 2A) and a single iteron (Figure 2B) dsDNA sequences (14–16) were tested in vitro by means of electrophoretic mobility shift assays (EMSA). Both indigoids completed the formation of RepA-operator complexes, with an apparent IC50 between 50–100 μM. In contrast, the tartrate ligand derivatives mentioned earlier worked in a much poorer way (IC 50–100 μM) with a single molecule bound per protein dimer (ligand-to-protein monomer stoichiometry, n = 0.69) (23), as predicted by docking (Figure 1B). This high affinity protein–ligand interaction is enthalpy-driven (ΔH = −9.73 kcal/mol; ΔS = −0.003 kcal/mol). The free energy of binding calculated from the ITC measurements (ΔG = −9.19 kcal/mol) is close to that estimated in the docking procedure (ΔG = −9.19 kcal/mol) (Figure 1B), supporting the reliability of the latter. A second transition became evident at the highest concentration range of S4-indigo tested, probably reflecting multiple and non-specific ligand binding (Kd = 23.3 μM, n = 1.43; approximately three molecules bound per WH1 dimer; ΔG = 0.3 kcal/mol), which results in the formation of aggregates. We cannot exclude that these non-specific binding sites in RepA-WH1 could become available due to some structural change in the protein induced by the primary S4-indigo binding event.

**Thermodynamics of S2/S4-indigoids binding to RepA-WH1**

Isothermal Titration Calorimetry (ITC; 23) was then used to get a reliable measure of the binding constants and overall energetics of the interactions between S2/S4 indigoids and RepA-WH1 dimers. The latter was used, rather than full length RepA, due to its higher solubility (16). The titrations performed with S4-indigo (Figure 3) were exothermic, pointing to a major site (Kd = 1/Ka = 0.23 μM) with a single molecule bound per protein dimer (ligand-to-protein monomer stoichiometry, n = 0.69) (23), as predicted by docking (Figure 1B). This high affinity protein–ligand interaction is enthalpy-driven (ΔH = −9.73 kcal/mol; ΔS = −0.003 kcal/mol). The free energy of binding calculated from the ITC measurements (ΔG = −9.19 kcal/mol) is close to that estimated in the docking procedure (ΔG = −9.19 kcal/mol) (Figure 1B), supporting the reliability of the latter. A second transition became evident at the highest concentration range of S4-indigo tested, probably reflecting multiple and non-specific ligand binding (Kd = 23.3 μM, n = 1.43; approximately three molecules bound per WH1 dimer; ΔG = 0.3 kcal/mol), which results in the formation of aggregates. We cannot exclude that these non-specific binding sites in RepA-WH1 could become available due to some structural change in the protein induced by the primary S4-indigo binding event.
Titration of WH1 with S2-indigo did not allow a reliable determination of either affinity constants or stoichiometry, due to non-specific binding which also gave way to protein aggregation.

S2/S4-indigos inhibit WH1-A31V assembly into amyloids

The ability of the S2/S4-indigo molecules to inhibit the dsDNA-induced assembly of the hyper-amyloidogenic WH1-A31V mutant into fibres was then checked. In summary, purified protein and dsDNA oligonucleotide were incubated in the presence of equimolar amounts, or 3- or 9-fold molar excess, of the indigoid ligands for a month, to allow for full maturation of the fibres (18). Samples were then inspected by means of transmission electron microscopy (TEM) after negative staining (Figure 4). In the control samples incubated in the absence of either of the indigoid stains, bundles of long, 75 nm thick, amyloid fibres (18), were evident (Figure 4A). However, samples incubated with S2-indigo (up to 3:1 ligand-to-protein ratio) showed smaller and thinner (ø = 25 nm) fibres. These became even shorter and curved at a higher molar ratio (9:1). When S4-indigo was the molecule tested, the result was very similar to that for the di-sulphonated molecule at the low molar ratios (Figure 4B), but at the highest (9:1) no fibre was evident, rendering instead an amorphous aggregate, similar to that reported (18) when WH1-A31V is incubated in the absence of dsDNA (Figure 4A). In conclusion, targeting with acidic indigo derivatives the binding site of effector dsDNA molecules in WH1-A31V is an efficient way to inhibit the assembly of the latter into amyloid fibres.

DISCUSSION

A central issue in current protein science is the survey of factors promoting amyloidogenesis. Natural and synthetic

Figure 3. Thermodynamic characterization (ITC) of the binding of S4-indigo to WH1-WT. (A) Binding isotherm for the titration of protein (98 μM) with the ligand (1.41 mM reservoir; concentration interval: 7.8–172 μM). (B) Integration of data plotted against the molar ratio between S4-ligand and WH1 (as monomers). The curve fits to a major binding site (#1) with a secondary site (#2) becoming evident at higher dye:protein ratios.

Figure 4. S2/S4-indigos inhibit DNA-promoted WH1-A31V amyloidogenesis according to TEM (magnification bars: 200 nm). (A) Control aggregates (top) and fibres (bottom) observed after 1 month incubation in the absence or presence of opsp dsDNA, respectively (18). (B) Samples including three different indigoid-to-protein (and dsDNA) molar ratios. S2 (top row) and, specially, S4-indigo (bottom row) lead to shorter and thinner fibres and, at the highest concentration tested, S4-indigo yields protein aggregates alike those found in the absence of DNA (see A).
polyanions, in particular RNA and DNA, have been found to enhance the assembly of the mammalian PrP into amyloids (3–11). Similarly, the dimeric WH1 domain of the bacterial plasmid replication protein RepA, besides its role in regulating replication initiation by the full length protein (13; Figure 5A), assembles into distinct amyloid aggregates, including spherulites and ordered fibres, upon transient binding to small specific dsDNA sequences (18) (Figure 5B). These are not constituents of the final aggregated particles which are mainly (if not exclusively) formed by RepA-WH1 (18), in agreement with the protein only nature of any amyloid structure (1,2), assembled through an amyloidogenic motif located opposite to the DNA-binding surface in WH1 structure (17).

We have found that di- and tetra-sulphonated derivatives of the indigo stain dock at the dsDNA recognition surface in RepA-WH1 (17), thus competing protein–DNA interactions and subsequently inhibiting the assembly of WH1 into amyloid fibres. The interaction of a single S4-indigo molecule with a WH1 dimer primarily occurs through an amyloidogenic motif located opposite to the DNA-binding surface in WH1 structure (17).

Figure 5. A scheme of the conformational transactions experienced by pPS10 RepA protein. (A) For the whole RepA protein, the balance between dimers (dRepA) and monomers (mRepA) is modulated by the target iteron DNA sequence (blue). WH2 binding implies dimer dissociation and a local secondary structure conversion (α → β) in WH1 (13,16), which becomes a metastable intermediate (the unbound domain depicted as a star) (14). This short-lived species is stabilized through WH1 binding to the iteron (15). On the contrary, dRepA can bind to the operator DNA sequence (red) without altering its association state and with minor effects (if any) on the structure of the WH1 domain (represented as a cube). Overall, the binding to the iteron sequence defines an irreversible refolding trajectory for WH1. (B) The extent of the remodelling exerted by the core of those target DNA sequences in the WH1 domain, in combination with the amyloidogenic enhancer A31V mutation, lead to the assembly of the protein into either ordered fibres, or spheroids, or large and irregular aggregates (18). Their formation can be interpreted as paralleling the pathway outlined in (A). As it is shown in this article, binding of sulphonated indigoids (S4 molecule depicted) to WH1 dimers competes with dsDNA binding, thus inhibiting amyloidogenesis and leading to small, grainy protein aggregates.
contribution of electrostatic interactions to $\Delta G$. However, the sulphonated indigoids can still exert their inhibitory action on WH1 amyloidogenesis during the prolonged incubations required for nucleation and growth of the protein fibres. The interaction of S4-indigo with WH1 dimers is specific, albeit at high ligand concentrations multiple molecules bind non-specifically. This would be related to the kind of large micellar complexes described for some hits coming out frequently in high throughput screenings of diverse protein targets (25).

Our results suggest that the promoting action of dsDNA on WH1 amyloidogenesis does not merely consist in shielding the electropositive patch in the protein (17) to allow for fibre assembly (18), since the indigoids would have the same effect and yet they do not yield ordered fibres, but increasingly disordered structures and, finally, amorphous aggregates. Thus, transient interactions with operator dsDNA should promote in WH1 some structural alteration that must be crucial for amyloidogenesis, contributing to the changes in the CD spectra previously attributed just to the assembly of the amyloidogenic stretch into a cross-$\beta$ spine (18).

Many small molecules, including a number of indol derivatives (26), have been described to directly interfere with protein aggregation through amyloidogenic stretches (19,20). However, to our knowledge those reported here are the first targeting a protein region not directly assembled into a cross-$\beta$ structure, but a site recognized by an allosteric effector of protein amyloidogenesis. The sulphonated indigoids, due to their specificity for the dsDNA binding pocket in RepA-WH1, are unlikely to have general applicability for other amyloidogenic proteins or even to bind to close RepA homologues, such as the dimers of the RepE protein, which do not bind to operator DNA through the WH1 domain (27). This is probably due to its reduced electropositive character, derived from the lack of basic residues equivalent to Arg81, 91 and 93 in RepA. The S2/S4-indigos also have the intrinsic limitations to membrane permeability derived from their acidic character. Notwithstanding, this work illustrates both the contribution of co-factors (dsDNA) to the assembly of protein amyloids and the feasibility of targeting their binding sites to design anti-amyloidogenic drugs.

ACKNOWLEDGEMENTS

We are indebted to the staff of the oligonucleotide synthesis and electron microscopy facilities at CIB for their technical support, to Dr Javier Ruiz for suggestions on ITC and to Drs Federico Gago and Antonio Romero for encouragement. Our groups are part of the Bioinformatics Integrative Platform for structureE-based Drug Discovery (BIPEDD-CM), funded by the Regional Government of Madrid (P-BIO-0214-2006). This work has been financed by grants of Spanish MEC (BMCC2003-00088 and BFU2006-00494) to R.G. Funding to pay the Open Access publication charges for this article was provided by grant BFU2006-00494.

Conflict of interest statement. None declared.

REFERENCES


