Published in final edited form as:

Biomacromolecules. 2012 July 9; 13(7): 1996–2002. doi:10.1021/bm300455f.

Designed Ankyrin Repeat Proteins as Scaffolds for Multivalent Recognition

Jessica J. Hollenbeck*, Derek J. Danner, Rachel M. Landgren, Thomas K. Rainbolt, and Danielle S. Roberts

Department of Chemistry Trinity University San Antonio, TX 78212

Abstract

Ankyrin repeat (AR) proteins are composed of tandem repeats of a basic structural motif of ca. 33 amino acid residues that form a β -turn followed by two antiparallel α -helices. Multiple repeats stack together in a modular fashion to form a scaffold that is ideally suited for the presentation of multiple functional groups and/or recognition elements. Here we describe a biosynthetic strategy that takes advantage of the modular nature of these proteins to generate multivalent ligands that are both chemically homogeneous and structurally well-defined. Glycosylated AR proteins cluster the tetrameric lectin concanavalin A (Con A) at a rate that is comparable to the rate of Con A aggregation mediated by globular protein conjugates and variable density linear polymers. Thus, AR proteins define a new class of multivalent ligand scaffolds that have significant potential application in the study and control of a variety of multivalent interactions.

Introduction

Multivalent binding is important in a variety of biological processes, including viral entry, cell surface adhesion, and fertilization, and it is involved in the self-assembly of supramolecular architectures.^{1–3} In recent years, substantial interest in this field has produced a wide range of synthetic multivalent ligands that present specific binding or recognition epitopes from a common scaffold. The scaffolds range in size and shape from polymeric materials and nanoparticles to small molecules with rigid cores including cyclodextrin, calixarene, and porphyrin conjugates.^{4–6}

One disadvantage of using small molecule scaffolds is that they are limited in the number of recognition epitopes that can extend from the central core and thus in the maximum distance between the recognition epitopes. Polymeric scaffolds, on the other hand, can present tens to hundreds of recognition epitopes, but they often suffer from some degree of heterogeneity in their molar mass distribution. Protein polymers containing repeating sequences of amino acids can overcome both of these limitations. They are chemically homogeneous; the DNA template that codes for their synthesis precisely defines their sequence, their stereochemistry, and to a large degree, their three-dimensional conformation, and theoretically, there is no limit to their size or to the number of recognition epitopes they can display.

*To whom correspondence should be addressed. Jessica J. Hollenbeck, telephone: (210) 999-7659; fax: (210) 999-7569; Jessica.Hollenbeck@Trinity.edu.

Supporting Information Amino acid sequence of ANK3a, SDS-polyacrylamide gel of glycosylated ANK3a C34, characterization of biotinylated ANK3a C14, ¹H and ¹³C NMR spectra of compounds 1 and 3. This information is available free of charge via the Internet at <http://pubs.acs.org>.

Monodisperse protein polymers, synthesized recombinantly in *E. coli*, can act as multivalent ligand scaffolds when chemically modified by recognition epitopes bearing reactive functional groups. This strategy has been used to create precisely defined macromolecules that self-assemble into liquid crystalline phases,^{7,8} and more recently, for the multivalent display of bioactive peptides and peptoids⁹ and carbohydrate epitopes.^{10–12} These protein polymers have either a random coil or an extended helical conformation; no other secondary structures or folded domains have been characterized. Given that α -helices are known to have helix to coil transitions at or below physiological temperatures, and that there can be a large entropic penalty associated with polymer organization upon binding,^{1,13–16} we sought to develop a monodisperse protein polymer with a well-defined three-dimensional conformation as a rigid scaffold to study the degree to which conformational flexibility and pre-organization contribute to multivalent binding. To that end, we describe here the design and synthesis of a multivalent ligand scaffold that is both chemically homogeneous and structurally well-defined.

Experimental Section

Materials

All reagents and solvents were purchased from commercially available sources and used as received without further purification. 9-fluorenylmethyl *N*-(2-hydroxyethyl)-carbamate (2-(Fmoc-amino)ethanol), avidin, boron trifluoride diethyl etherate, concanavalin A, 4'-hydroxyazobenzene-2-carboxylic acid (HABA), and α -D-mannose pentaacetate were purchased from Sigma Aldrich; 3-maleimidopropionic acid N-hydroxysuccinimide ester was purchased from Obiter Research LLC, and maleimide-PEG₂-biotin and immobilized TCEP disulfide reducing gel were purchased from ThermoScientific.

Characterization

Synthetic reactions were monitored by thin layer chromatography (TLC) carried out on TLC plates pre-coated with silica gel (200 μ m thick) visualized by using UV (254 nm) or charring with at least one of the following solutions: a *p*-anisaldehyde stain (3.7 mL *p*-anisaldehyde, 1.5 mL glacial acetic acid, 5 mL of concentrated H₂SO₄, 135 mL absolute ethanol) or ninhydrin stain (1.5 g ninhydrin, 100 mL *n*-butanol, 3.0 mL glacial acetic acid). Flash chromatography was performed on silica gel (40–63 μ m, 60 Å pore size) using ACS grade ethyl acetate, hexanes, methanol, and/or dichloromethane as the eluent. ¹H and ¹³C NMR spectra were collected on Varian NMR spectrometers. Chemical shifts are reported in parts per million relative to residual solvent peaks (CHCl₃: ¹H: δ 7.26, ¹³C: δ 77.0; D₂O: ¹H: δ 4.80).

2'-(Fmoc-amino)ethyl 2,3,4,6-tetra-O-acetyl- α -D-mannopyranoside (1)—Boron trifluoride diethyl etherate (5.5 mL, 44.6 mmol) was added dropwise to a solution of α -D-mannose pentaacetate (5.00 g, 12.8 mmol) and 2-(Fmoc-amino)ethanol (4.00 g, 14.1 mmol) in 50 mL of CH₂Cl₂ under a nitrogen atmosphere at 0 °C. After a few minutes, the mixture was heated to reflux and allowed to stir for 36 hours. The reaction was then quenched with water and extracted with CH₂Cl₂. The CH₂Cl₂ layer was collected, dried with MgSO₄, and concentrated *in vacuo*. The resulting residue was purified by silica gel chromatography using 1:1 hexanes : ethyl acetate as the eluent. Yield 4.08 g (52%). ¹H NMR (300 MHz, CDCl₃): δ 2.01 (s, 3H, OAc), 2.03 (s, 3H, OAc), 2.08 (s, 3H, OAc), 2.17 (s, 3H, OAc), 3.37 – 3.64 (m, 2H, CH₂NH), 3.73 – 3.84 (m, 1H, OCH₂CH₂), 3.93 – 4.04 (m, 1H, OCH₂CH₂), 4.06 – 4.16 (m, 2H, H-5, H-6), 4.21 – 4.29 (m, 2H, H-6, OCH₂CH), 4.42 (d, J = 7.0 Hz, 2H, OCH₂CH), 4.83 (s, 1H, H-1), 5.20 (m, 1H, NH), 5.24 – 5.37 (m, 3H, H-2, H-3, H-4), 7.29 – 7.35 (m, 2H, ArH), 7.38 – 7.43 (m, 2H, ArH), 7.61 (d, J = 7.4 Hz, 2H, ArH), 7.77 (d, J = 7.3 Hz, 2H, ArH). ¹³C NMR (126 MHz, CDCl₃): δ 20.59, 20.62, 20.79, 40.59, 47.18, 62.47,

66.07, 66.71, 67.54, 68.66, 69.00, 69.36, 97.68, 119.93, 125.05, 127.02, 127.65, 141.22, 143.84, 143.91, 156.42, 169.64, 169.91, 169.96, 170.53. MS (HR-ESI): $[M + H]^+$ calcd for $C_{31}H_{36}NO_{12}$ 614.2238, found 614.2252.

2'-Aminoethyl α -D-mannopyranoside (2)—2'-(Fmoc-amino)ethyl 2,3,4,6-tetra-*O*-acetyl α -D-mannopyranoside (**1**) (1.01 g, 1.6 mmol) was stirred in methanol (58 mL) with sodium methoxide (2.4 mL of a 0.5 M solution in methanol) at room temperature overnight. The reaction mixture was concentrated *in vacuo*; the resulting residue was dissolved in water, and the aqueous solution was extracted with CH_2Cl_2 . The aqueous layer was lyophilized, and the product used without further purification.

3'-Maleimidopropanamidoethyl α -D-mannopyranoside (3)—3-maleimidopropionic acid N-hydroxysuccinimide ester (639 mg, 2.4 mmol) was added to a solution of the lyophilized solid (**2**) resuspended in anhydrous methanol (20 mL). The reaction mixture was stirred under a nitrogen atmosphere at room temperature for 2 hours. The solvent was removed *in vacuo*, and the final product was purified by silica gel chromatography using a gradient of methanol in CH_2Cl_2 as the eluent. Yield 252 mg (44% over 2 steps). 1H NMR (500 MHz, D_2O) δ 2.53 (t, $J = 7.3$ Hz, 2H, CH_2CH_2N), 3.29–3.44 (m, 2H, OCH_2CH_2NH), 3.52–3.67 (m, 3H, OCH_2CH_2NH , H-5), 3.72–3.82 (m, 5H, H-3, H-4, H-6, CH_2CH_2N), 3.88 (dd, $J = 12.3, 2.1$ Hz, 1H, H-6), 3.92 (dd, $J = 3.4, 1.7$ Hz, 1H, H-2), 4.84 (d, $J = 1.7$ Hz, 1H, H-1), 6.88 (s, 2H, $CH=CH$). ^{13}C NMR (126 MHz, D_2O): δ 34.39, 34.57, 38.84, 60.78, 65.63, 66.59, 69.95, 70.37, 72.67, 99.58, 134.38, 172.53, 173.48. MS (HR-ESI): $[M + Na]^+$ calcd for $C_{15}H_{22}N_2O_9Na$ 397.1223, found 397.1228.

Plasmid Construction—Plasmid pANK3a, encoding the designed ankyrin repeat protein E3_5, was constructed from vector pHKB26E3_5 (kindly provided by A. Plückthun).¹⁷ The coding sequence for E3_5 was amplified by the polymerase chain reaction and cloned into a pET-30 Ek/LIC vector using ligation independent cloning (EMD Biosciences). The resulting plasmid, pANK3a, contains the coding sequence for the 154 amino acid protein E3_5, preceded by His₆ and S tags (see Supporting Information figure S1). Plasmids coding for the cysteine-containing variants ANK3a C1 (N33C), ANK3a C14 (N33C K132C), and ANK3a C34 (N99C K132C) were constructed from pANK3a by site-directed mutagenesis using either a QuikChange® II Site-Directed Mutagenesis Kit or a QuikChange® Multi Site-Directed Mutagenesis Kit (Agilent). The primary sequence of each construct was confirmed by DNA sequencing.

Protein Expression and Purification—Proteins were overexpressed in *Escherichia coli* strain BL21 (DE3) pLysS using the T7 expression system.¹⁸ Cells derived from a single colony were grown in Luria Broth containing 30 μ g/mL of kanamycin at 37 °C. Protein expression was induced at $A_{600} = 0.5$ by addition of isopropyl-1-thio- β -D-galactopyranoside (IPTG) to a final concentration of 1 mM. Cells were harvested after 4 hours by centrifugation and lysed by freezing and sonication in 50 mM sodium phosphate, 300 mM NaCl, 10 mM imidazole, pH 8.0. After sonication, the cells were treated with lysozyme (1 mg/5 mL of cell lysate) and DNase I (2.5 mg in 1 mL of 10 mM Tris-Cl (pH 7.5), 150 mM NaCl, 1 mM $MgCl_2$), then centrifuged. 2-Mercaptoethanol was added to the supernatant to a final concentration of 20 mM, and the lysate was applied to a Ni-nitrilotriacetic acid column (4 mL column volume) according to the manufacturer's instructions (QIAGEN). The protein concentration was determined by UV absorbance in 6 M guanidinium hydrochloride and 0.02 M phosphate (pH 6.5) assuming an extinction coefficient of 1470 $M^{-1} cm^{-1}$ for Tyr at 275 nm.¹⁹ Approximately 3 mg of purified protein were isolated per liter of cell culture.

Chemical Conjugation—Cysteine-containing proteins were dialyzed into conjugation buffer (50 mM sodium phosphate, 300 mM NaCl, 5 mM EDTA, pH 7.0) and reduced by incubation with immobilized tris(2-carboxyethyl)phosphine (TCEP) disulfide reducing gel on a rocker platform for 2 hours at room temperature. The protein was recovered by centrifugation using a spin-cup column, then incubated with an excess of maleimide-PEG₂-biotin or maleimide-activated mannose for 2 hours at room temperature. The incubation was followed by extensive dialysis over a minimum of 16 hours, and the functionalized proteins were analyzed by ESI mass spectrometry at the Institutional Mass Spectrometry Laboratory at UTHSC San Antonio.

CD Spectroscopy—CD spectra were acquired with a Jasco J-815 spectropolarimeter. Samples were prepared in 50 mM sodium phosphate, 300 mM NaCl, 1 mM DTT, pH 8.0. The wavelength dependence of $[\theta]$ was monitored at 25 °C in 1 nm increments. Five scans of each sample were recorded, and the average signal plotted as a function of wavelength. The temperature dependence of the signal at 222 nm was monitored from 4 to 95 °C in 0.5 °C increments for each protein at a concentration of 10 μ M in 50 mM sodium phosphate, 300 mM NaCl, 1 mM DTT, pH 8.0.

HABA Displacement Assay—This assay was adapted from one described previously.²⁰ 900 μ L of a solution containing 300 μ M 4'-hydroxyazobenzene-2-carboxylic acid (HABA) and 0.5 mg/mL avidin in 100 mM sodium phosphate, 150 mM sodium chloride, pH 7.2 were added to a disposable microcuvette with a 1 cm pathlength. Absorbance data were recorded from 400 – 600 nm. 100 μ L of biotinylated ANK3a C14 were added, and the displacement of HABA was monitored by the change in absorbance at 500 nm.

Dynamic Light Scattering (DLS)—A stock solution of Con A was prepared by dissolving the lectin in 10 mM Tris-Cl, pH 7.5, 150 mM NaCl, 1 mM CaCl₂, and 1 mM MnCl₂. The resulting solution was mixed and filtered (2 μ m). The concentration of the filtered solution was determined by UV absorbance ($E^{1\%}$ at 280 nm = 12.4). Glycosylated ANK3a C14 and Con A were mixed together and DLS measurements were conducted immediately using a Delsa Nano C from Beckman Coulter, Inc. (Fullerton, CA) equipped with a laser diode operating at 658 nm. Scattered light was detected at a 15° angle and analyzed using a log correlator over 30 accumulations. The photomultiplier aperture and the attenuator were automatically adjusted to obtain a photon counting rate of ca. 10 kcps.

Kinetic Light Scattering—Turbidity measurements were made using Con A at a final concentration of 5 μ M tetramer in 10 mM Tris-Cl, pH 7.5, 150 mM NaCl, 1 mM CaCl₂, and 1 mM MnCl₂. 500 μ L were added to a dry quartz microcuvette with a 1 cm pathlength. Glycosylated ANK3a C14 was added; the solution was mixed vigorously for 5 seconds using a micropipette and then placed in the spectrometer. Absorbance data were recorded at 420 nm for 10 min at 1 Hz. The resulting solution was allowed to stand at room temperature for 1 hour. 10 μ L of a 50 mM solution of methyl- α -D-mannopyranoside were added; the solution was mixed vigorously for 5 seconds, and the change in absorbance at 420 nm was monitored again as a function of time.

Results and Discussion

Repeat proteins are protein polymers that contain tandem repeats of a simple structural motif.²¹ These repeating motifs are present in approximately 14% of all proteins, and they mediate numerous protein-protein interactions both inside and outside the cell.²² There are numerous types of repeats including tetratricopeptide, ankyrin, armadillo/HEAT, and leucine-rich repeats.

The ankyrin repeat (AR) is one of the most common protein sequence motifs.²³ AR proteins are found in all species and are fundamental to biological processes ranging from transcriptional regulation to cytoskeletal organization.²⁴ Each repeat consists of ca. 33 amino acids that fold into a β -turn followed by two antiparallel α -helices. The number of repeats within a single domain can vary from as few as three to more than twenty;^{25,26} these repeats stack together in a modular fashion to form elongated structures of varying size depending on the number of repeats. Plückthun and coworkers have demonstrated that consensus designed AR proteins can be expressed in high yields in bacteria as soluble, monomeric proteins that are stable at temperatures $>85^{\circ}\text{C}$.¹⁷ Given this extraordinary stability and the fact that the number of repeats, and thus the valency of the scaffold, can be varied systematically, we hypothesized that designed AR proteins would be ideally suited for the presentation of multiple functional groups and/or recognition elements in a structurally well-defined three-dimensional array.

The designed protein ANK3a has three consensus repeats flanked by N- and C-terminal capping repeats. AR proteins can sustain a considerable variety of amino acid substitutions,^{24,27–30} and in particular, there is little consensus at the positions between the Asp and Gly that define the β -turn in each repeat. These residues extend away from the hydrophobic core of the folded domain (Figure 1) and are often involved in the recognition of specific ligands by naturally-occurring AR proteins.^{17,25} We envisaged that residues at these positions would be amenable to post-translational modification with small molecule recognition epitopes. As an initial demonstration of our approach, we modified the β -turns at the two ends of the protein.

Reactive thiols can be introduced into the protein polymer at defined positions by site-directed mutagenesis, and in aqueous solution at pH 7, they can be modified specifically by recognition epitopes bearing a maleimide functional group.³¹ We introduced two cysteines into the coding sequence for ANK3a. The N33C K132C variant (ANK3a C14) contains one cysteine in the N-terminal (first) repeat and a second cysteine in the C-terminal capping repeat. ANK3a C14 has a CD spectrum that is characteristic of a highly helical protein (Figure 2a), and its thermal stability is comparable to the consensus designed AR protein E3_5 (Figure 2b).¹⁷

While use of a monodisperse protein polymer limits one source of polydispersity in the synthesis of a multivalent ligand, the efficiency of the coupling reaction between the reactive groups on the scaffold and the pendant moieties can introduce another. We used commercially-available biotin-PEG₂-maleimide to identify reaction conditions that would yield complete conjugation of the pendant thiols in ANK3a C14. ESI mass spectrometry of the protein conjugate resulting from treatment of the AR scaffold with a four-fold molar excess of biotin-PEG₂-maleimide shows near complete conjugation (see Supporting Information figure S3). Importantly, biotinylation had little effect on the overall secondary structure or thermal stability of the AR domain (Figure 2).

Biotinylated ANK3a C14 binds to avidin (see Supporting Information figure S4) demonstrating that the AR scaffold does not interfere with the accessibility or biological function of the pendant recognition epitopes. Encouraged by these results, we sought to determine whether *multiple* ligands displayed from the same AR scaffold could bind simultaneously to a single protein target (i.e., in a multivalent fashion).

Multivalent interactions can occur by a variety of different mechanisms. Multiple ligands on a single scaffold can bind simultaneously to multiple sites within a single oligomeric protein target; multiple ligands on a single scaffold can cluster multiple protein targets, or the increased effective concentration of a multivalent ligand can give rise to increased apparent

affinity.^{2,32} The jack bean lectin concanavalin A (Con A) binds to α -linked manno- or glucopyranosides. A number of multivalent ligands for Con A have been described,^{1,4,32,33} making this an ideal model system with which to test the ability of AR proteins to serve as multivalent ligand scaffolds.

Con A exists as a homotetramer at neutral pH with the binding sites in each monomer separated by 65 Å.³⁴ From the crystal structure of the designed protein E3_5,³⁵ we estimate that the distance between the two sugars on the ANK3a C14 scaffold is 25 Å (Figure 1a) precluding the simultaneous binding of glycosylated ANK3a C14 to two sites within the same Con A tetramer. The glycoconjugate should, however, be able to effectively crosslink (or cluster) multiple Con A tetramers.

We synthesized a maleimide-activated mannose derivative by taking advantage of well-established methodology for the synthesis of 2'-aminoethyl glycosides.^{36–38} Specifically, Lewis acid mediated glycosylation of mannose pentaacetate with Fmoc-protected aminoethanol gave compound **1**.³⁹ The Fmoc and acyl protecting groups were removed simultaneously to give intermediate **2**, which was then coupled to 3-maleimidopropionic acid *N*-hydroxysuccinimide ester (Scheme 1). Fully reduced ANK3a C14 was treated with four molar equivalents of compound **3**, and after exhaustive dialysis, ESI mass spectrometry confirmed complete glycosylation of the ANK3a C14 scaffold (Figure 3).

The binding of multivalent carbohydrates to Con A results in the formation of multi-dimensional cross-linked complexes.^{40–43} These complexes can be detected by dynamic light scattering (DLS); in fact, at high mannose concentrations, they rapidly precipitate from solution. We hypothesized that multivalent AR proteins (bearing 2 mannose residues) would effectively crosslink Con A tetramers forming a supramolecular complex that would initiate the aggregation of additional Con A oligomers.

Free Con A has a peak diameter of 7 nm (data not shown). When Con A was mixed with ANK3a (no mannose), we observed an average particle size between 50–100 nm that we attribute to nonspecific binding (Figure 4). Importantly, there was little to no change in the average particle size over time. Similar results were observed with glycosylated ANK3a C1 (bearing only a single mannose ligand) and with glycosylated ANK3a C14 at sub-stoichiometric ratios of mannose to Con A monomer. However, with glycosylated ANK3a C14, as the concentration of mannose increased with respect to the concentration of Con A monomers, we observed the formation of larger aggregates that increased in size over time (Figure 4). These aggregates result from extensive crosslinking of Con A tetramers. At molar ratios greater than 4:1 mannose : Con A monomer, precipitation was visible in the sample upon mixing, precluding analysis by DLS. These results suggest that Con A aggregation is dependent on the simultaneous binding of two Con A tetramers to a single ANK3a C14 scaffold, thus demonstrating a multivalent interaction. They do not, however, demonstrate any cooperativity in binding or increased avidity due to multivalency.

To further establish that the observed aggregation was due to specific protein – carbohydrate interactions and **not** due to nonspecific protein aggregation, we used a kinetic light scattering assay in which the precipitation of Con A is monitored by a change in the visible region of the electromagnetic spectrum (Figure 5). If precipitate formation is mediated by specific interactions involving the pendant sugars, then the addition of a competitive ligand should reverse the aggregation process. Indeed, upon the addition of 1 mM methyl- α -D-mannopyranoside, we observed rapid dissolution of the aggregates. Taken together, these results demonstrate that AR proteins are viable scaffolds for the multivalent display of small molecule recognition epitopes.

Over the past ten years, ankyrin and other repeat proteins have been successfully designed for applications in viral therapy, tumor targeting, and drug discovery.^{44,45} In the majority of these applications, the repeat protein has served as a bispecific or bifunctional scaffold. Recently, in fact, Simon *et al.* have synthesized a bifunctional AR protein by introducing a clickable methionine surrogate, azidohomoalanine (Aha), at the N-terminus of the protein domain and a single cysteine at the C-terminus.⁴⁶ By comparison, our design has several advantages. For example, (1) the synthesis of the scaffold does not require the incorporation of a non-natural amino acid; (2) the valency of the scaffold can be varied systematically by changing the number of repeats or the number of cysteines incorporated in the AR domain, and (3) the distance between the functional groups or recognition epitopes can be controlled precisely by varying the position of the cysteine-containing repeats.

To illustrate the third point, we synthesized an additional AR protein, a N99C K132C variant (ANK3a C34) that contains cysteines in the third and fourth repeats of the AR domain. Upon glycosylation, the ANK3a C34 scaffold displays two mannose residues separated by approximately 8.5 Å.³⁵ Using the kinetic light scattering assay described above, we observed an initial rate (k_i) of Con A aggregation mediated by glycosylated ANK3a C14 of 0.045 ± 0.004 AU/min. Under the same conditions, the initial rate of Con A aggregation mediated by glycosylated ANK3a C34 was 0.18 ± 0.01 AU/min, four times greater than what was observed with the C14 glycoconjugate (Figure 5). The extent of aggregation was also greater with glycosylated ANK3a C34, consistent with the observation that decreased distance between mannose residues in bivalent glycoconjugates correlates with increased yield of cross-linked Con A complexes.^{47,48} These results demonstrate that the distance (spacing) between recognition epitopes on the AR scaffold can have a dramatic effect on its function.

The rates observed for Con A aggregation by glycosylated AR proteins compare well with those observed using generation 1 glycodendrimers with 8 terminal mannose residues, polydisperse polyethylene-maleic anhydride polymers with 200 mannose residues per polymer, and variable density linear polymers generated by ring-opening metathesis polymerization.^{32,49} Interestingly, dimers and trimers of mannose as well as globular proteins bearing two mannose residues per scaffold, failed to aggregate Con A in the same assay.³² These results, in the context of the fact that the ANK3a C14 and C34 scaffolds produce only bivalent displays of mannose, suggest that AR proteins have significant potential to mediate multivalent interactions when protein aggregation and receptor clustering are involved.

Conclusion

To date, multivalent ligand scaffolds with vastly different architectures have been described including globular protein conjugates,^{50,51} synthetic dendrimers, and linear polymers.³² Ankyrin repeat proteins define a new class of multivalent ligand scaffolds. Given their modular nature, AR proteins with different numbers of repeats are readily accessible, and our biosynthetic strategy allows the display a variety of different recognition epitopes. Indeed, their ease of synthesis, their solubility, and their remarkable stability, make designed AR proteins attractive scaffolds for the study and control of multivalent interactions in biotechnology and in materials applications.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors acknowledge D. Kelly, M. McManus, D. Samples, D. Samples, and C. Wiley for preliminary contributions to this work. We thank Prof. Adam Urbach (Trinity University) for helpful discussions, Prof. Andreas Plückthun (Universität Zürich) for pHKB26E3_5, Dr. Nam Lee and Prof. Karen Wooley (Texas A&M University) for help with dynamic light scattering measurements, and Mr. Kevin Hakala and Prof. Sue Weintraub (UTHSCSA) for assistance with mass spectrometry. This work was supported by grants from the Camille & Henry Dreyfus Foundation, Research Corporation (CC10655), and the National Science Foundation (DBI-0718766 and CHE-0957839). Summer research fellowships to DJD, RML, TKR, and DSR were provided by the Welch Foundation, Prof. John A. Burke, and the Howard Hughes Medical Institute at Trinity University. The UTHSCSA Institutional Mass Spectrometry Laboratory is supported by NIH grant P30 CA54174. Molecular graphics images were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR001081).

References

1. Mammen M, Choi S-K, Whitesides GM. *Angew. Chem. Int. Ed.* 1998; 37:2755–2794.
2. Kiessling LL, Gestwicki JE, Strong LE. *Angew. Chem. Int. Ed.* 2006; 45:2348–2368.
3. Martos V, Castreño P, Valero J, de Mendoza J. *Curr. Opin. Chem. Biol.* 2008; 12:698–706. [PubMed: 18801458]
4. Choi, S-K. *Synthetic Multivalent Molecules*. John Wiley & Sons, Inc.; New Jersey: 2003.
5. Chabre YM, Roy R. *Adv. Carbohydr. Chem. Biochem.* 2010; 63:165–393. [PubMed: 20381707]
6. Deniaud D, Julianne K, Gouin SG. *Org. Biomol. Chem.* 2011; 9:966–979. [PubMed: 21173976]
7. Yu SM, Conticello VP, Zhang G, Kayser C, Fournier MJ, Mason TL, Tirrell DA. *Nature.* 1997; 389:167–170. [PubMed: 9296493]
8. Zhuravel MA, Davis NE, Nguyen ST, Koltover I. *J. Am. Chem. Soc.* 2004; 126:9882–9883. [PubMed: 15303837]
9. Davis NE, Karfeld-Sulzer LS, Ding S, Barron AE. *Biomacromolecules.* 2009; 10:1125–1134. [PubMed: 19361214]
10. Wang Y, Kiick K. *J. Am. Chem. Soc.* 2005; 127:16392–16393. [PubMed: 16305215]
11. Farmer RS, Kiick K. *Biomacromolecules.* 2005; 6:1531–1539. [PubMed: 15877375]
12. Maheshwari R, Liu S, Polizzotti BD, Wang Y, Kiick KL. *ACS Symposium Series.* 2008; 990:288–305. (Chemical Glycobiology).
13. Jencks, WP. *Catalysis in Chemistry and Enzymology*. Dover Publications, Inc.; New York: 1987.
14. Lundquist JJ, Toone EJ. *Chem. Rev.* 2002; 102:555–578. [PubMed: 11841254]
15. Reczek JJ, Kennedy AA, Halbert BT, Urbach AR. *J. Am. Chem. Soc.* 2009; 131:2408–2415. [PubMed: 19199617]
16. Nangreave J, Yan H, Liu Y. *J. Am. Chem. Soc.* 2011; 133:4490–4497. [PubMed: 21381740]
17. Binz HK, Stumpp MT, Forrer P, Amstutz P, Plückthun A. *J. Mol. Biol.* 2003; 332:489–503. [PubMed: 12948497]
18. Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW. *Methods Enzymol.* 1990; 185:60–89. [PubMed: 2199796]
19. Edelhoch H. *Biochemistry.* 1967; 6:1948–1954. [PubMed: 6049437]
20. Greene NM. *Methods Enzymol.* 1970; 18:418–424.
21. Cortajarena AL, Mochrie SGJ, Regan L. *Protein Sci.* 2011; 20:1042–1047. [PubMed: 21495096]
22. Andrade MA, Perez-Iratxeta C, Ponting CP. *J. Struct. Biol.* 2001; 134:117–131. [PubMed: 11551174]
23. Sedgwick SG, Smerdon SJ. *Trends. Biochem. Sci.* 1999; 24:311–316. [PubMed: 10431175]
24. Bork P. *Proteins: Struct. Funct. Genet.* 1993; 17:363–374. [PubMed: 8108379]
25. Gaudet R. *Mol. Biosyst.* 2008; 4:372–379. [PubMed: 18414734]
26. Walker RG, Willingham AT, Zuker CS. *Science.* 2000; 287:2229–2234. [PubMed: 10744543]
27. Michaely P, Bennett V. *Trends Cell Biol.* 1992; 2:127–129. [PubMed: 14731966]
28. Zweifel ME, Leahy DJ, Hughson FM, Barrick D. *Protein Sci.* 2003; 12:2622–2632. [PubMed: 14573873]

29. Lowe AR, Itzhaki LS. J. Mol. Biol. 2007; 365:1245–1255. [PubMed: 17113103]
30. Tang KS, Guralnick BJ, Wang WK, Fersht AR, Itzhaki LS. J. Mol. Biol. 1999; 285:1869–1886. [PubMed: 9917418]
31. Hermanson, G. Bioconjugate Techniques. Academic Press, Elsevier Science (USA); 1996.
32. Gestwicki JE, Cairo CW, Strong LE, Oetjen KA, Kiessling LL. J. Am. Chem. Soc. 2002; 124:14922–14933. [PubMed: 12475334]
33. Pieters R. Org. Biomol. Chem. 2009; 7:2013–2025. [PubMed: 19421435]
34. Derewenda Z, Yariv J, Helliwell JR, Kalb AJ, Dodson EJ, Papiz MZ, Wan T, Campbell J. EMBO J. 1989; 8:2189–2193. [PubMed: 2792084]
35. Kohl A, Binz HK, Forrer P, Stumpp MT, Plückthun A, Grütter MG. Proc. Natl. Acad. Sci. USA. 2003; 10:1700–1705. [PubMed: 12566564]
36. Hayes W, Osborn HMI, Osborne SD, Rastall RA, Romagnoli B. Tetrahedron. 2003; 9:7983–7996.
37. Ni J, Singh S, Wang L-X. Bioconjugate Chem. 2003; 14:232–238.
38. Sardžik R, Noble GT, Weissenborn MJ, Martin A, Webb SJ, Flitsch SL. Beilstein J. Org. Chem. 2010; 6:699–703. [PubMed: 20978609]
39. Han Z, Pinkner JS, Ford B, Obermann R, Nolan W, Wildman SA, Hobbs D, Ellenberger T, Cusumano CK, Hultgren SJ, Janetka JW. J. Med. Chem. 2010; 53:4779–4792. [PubMed: 20507142]
40. Bhattacharyya L, Khan MI, Brewer CF. Biochemistry. 1988; 27:8762–8767. [PubMed: 3242606]
41. Brewer CF. Chemtracts Biochem. Mol. Biol. 1996; 6:165–179.
42. Olsen LR, Dessen A, Gupta D, Sabesan S, Sacchettini JC, Brewer CF. Biochemistry. 1997; 36:15073–15080. [PubMed: 9398234]
43. Dam TK, Brewer CF. Methods Enzymol. 2003; 362:455–486. [PubMed: 12968382]
44. Boersma YL, Plückthun A. Curr. Opin. Biotechnol. 2011; 22:849–857. [PubMed: 21715155]
45. Gill DS, Damle NK. Curr. Opin. Biotechnol. 2006; 17:653–658. [PubMed: 17055245]
46. Simon M, Zangemeister-Wittke U, Plückthun A. Bioconjugate Chem. 2012; 23:279–286.
47. Dam TK, Roy R, Das SK, Oscarson S, Brewer CF. J. Biol. Chem. 2000; 275:14223–14230. [PubMed: 10799500]
48. Dam TK, Oscarson S, Roy R, Das SK, Page D, Macaluso F, Brewer CF. J. Biol. Chem. 2005; 280:8640–8646. [PubMed: 15632152]
49. Cairo CW, Gestwicki JE, Kanai M, Kiessling LL. J. Am. Chem. Soc. 2002; 124:1615–1619. [PubMed: 11853434]
50. Stowell CP, Lee YC. Adv. Carbohydr. Chem. Biochem. 1980; 37:225–281. [PubMed: 6996450]
51. Artner LM, Merkel L, Bohlke N, Beceren-Braun F, Weise C, Dervede J, Budisa N, Hackenberger CPR. Chem. Commun. 2012; 48:522–524.
52. <http://www.cgl.ucsf.edu/chimera>
53. Strupat K. Methods Enzymol. 2005; 405:1–36. [PubMed: 16413308]

Figure 1.

(a) Designed ankyrin repeat proteins can act as multivalent scaffolds for the display of a variety of recognition epitopes including biotin and α -methyl mannose. Shown here is an ankyrin repeat protein with three internal repeats (ANK3a) containing two cysteines as points of attachment; (b) side view of ANK3a demonstrating how the recognition epitopes extend away from the hydrophobic core of the protein; the protein structures shown in (a) and (b) were rendered in Chimera⁵² using coordinates from PDB ID: 1MJ0.³⁵

Figure 2.

(a) CD spectra of the designed ankyrin repeat protein ANK3a C14 (containing cysteines in the first and fourth repeats of the AR domain) before and after conjugation to biotin or mannose. The spectra were collected in 50 mM sodium phosphate, 300 mM NaCl, 1 mM DTT, pH 8.0. (b) Temperature dependence of the CD signal at 222 nm of 10 μM solutions in 50 mM sodium phosphate, 300 mM NaCl, 1 mM DTT, pH 8.0.

Figure 3.

ESI mass spectrum showing complete glycosylation (+ 374.13 Da / mannose conjugate) of an ANK3a derivative with two cysteine residues (isotopically averaged molecular weight: 21274.5). Expected mass of the bivalent conjugate = 22022.76 Da. (a) charge state envelope with m/z peaks corresponding to 19 different protonation states of the bivalent conjugate. Three charge states are assigned; (b) deconvoluted mass spectrum calculated from the charge state envelope shown in (a).⁵³

Figure 4.

Cumulants analysis of AR : ConA aggregates observed by dynamic light scattering. Average size of Con A aggregates upon mixing with glycosylated AR proteins at different molar ratios. The x-axis represents five consecutive data sets each consisting of 30 measurements. Error bars are $\pm$ S.D. The average time per data set was 2.4 minutes. Data with glycosylated ANK3a C14 (containing two mannose residues) are shown in red; glycosylated ANK3a C1 (containing a single mannose residue) ($\circ$), and ANK3a (unglycosylated) ($\bullet$). The concentration of Con A tetramer was $1.9 \mu\text{M}$. Concentrations of the AR protein conjugates were as follows: ($\bullet$) $15.4 \mu\text{M}$; ($\blacklozenge$) $7.7 \mu\text{M}$; ($\blacktriangle$) $3.9 \mu\text{M}$; ($\bullet$) $15.4 \mu\text{M}$; ($\circ$) $30.8 \mu\text{M}$.

Figure 5.

Con A aggregation was monitored at 420 nm for 10 min. The curves shown represent the average of three independent experiments with 5 μ M Con A tetramer and (●) 25 μ M glycosylated ANK3a C14 (50 μ M mannose) or (◻) 25 μ M glycosylated ANK3a C34 (50 μ M mannose). The initial rate of precipitation was determined by a linear fit to the steepest portion of the curve. After one hour, methyl- α -D-mannose was added to a final concentration of 1 mM, and dissolution of the aggregates was monitored at the same wavelength. Con A precipitation is dependent on the quaternary structure of the lectin. No aggregation was observed with succinylated Con A (○) 5 μ M dimer or (◻) 10 μ M dimer using the same concentration of glycosylated ANK3a C14.

Scheme 1.

Synthesis of a maleimide-activated mannose derivative for conjugation to cysteine-containing ankyrin repeat proteins.