NJC

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/njc

New Journal of Chemistry

RSC**Publishing**

ARTICLE

Helical Arylamide Foldamers: Structure Prediction by Molecular Dynamics Simulations

Zhiwei Liu, Ara M. Abramyan and Vojislava Pophristic[*]

Abstract We present a molecular dynamics (MD) study on a series of helical arylamide oligomers with systematically varying building blocks and linkage types. This study showcases a computational approach for the prediction of secondary structure properties of arylamide foldamers and their solution dynamics. We demonstrate that conformational characteristics of foldamers, such as number of units per turn, helical pitch, and pore diameter, can be predicted by MD simulations of small oligomers significantly shorter than the foldamers in question. Importantly, the curvature angle, the key geometrical parameter in helical arylamide structures, can be accurately determined by MD simulation of tetramers, entities with often less than one helical turn. The curvature angle is found to be a local property associated with one single residue/unit, which enables highly accurate predictive power for designing oligomers with various scaffolds and sizes. In addition, MD simulations with the improved force field parameters capture solvent effects in terms of both protic solvent competition with intramolecular H-bonds and solvophobic effects. The computational approach can provide useful insight into dynamical, mechanistic and functional properties of the arylamide oligomer class, which will facilitate rational design of foldamers.

1. Introduction

Over the past decade, the interest in design and application of foldamers, synthetic oligomers that adopt defined secondary structures in solution, has increased rapidly.[1,2] Aromatic oligoamide (arylamide) foldamers attracted attention in particular, because of a number of attractive properties, such as high stability and predictability of their folded conformations, tendency to crystallize and easiness to synthesys.[3] Consequently, arylamide foldamers have been used as potential anticancer therapeutics,[4,5] in molecular recognition,[6,7] catalysis[8] and nanostructures.[9]

As the name indicates, arylamide foldamer backbone comprises alternating aromatic and peptide groups, providing the basis for conformational control by hydrogen bonding (H-bonding), π−π stacking, and geometrical constraints. An important feature of the arylamide class of foldamers is their structure tunability, which stems from the possibility to change aromatic entities and/or positions of H-bonding substituents and thus influence the geometrical pattern and strength of H-bonding. This allows, for example, rational syntheses of molecular capsules with varying diameters of the central cavity for capturing diverse ligands,[6,7] design of molecular tweezers with a desired degree of flexibility,[10] or aromatic helices with a preferred level of stability.[11] Despite the established level of control, it

is still difficult to predict the interplay of the interactions that eventually determine the conformation of foldamers prior to conducting experimental studies.[12] It is because of this that the prominent researcher in the foldamer field, Samuel H. Gellman, dedicated his seminal paper[13] to the "daunting step" in the design of functional foldamers, which is understanding the "relationship between repetitive features of monomer structure and conformational properties at the polymer level".

A suitable computational approach could provide us with power to predict shapes and sometimes subtle but function-determining differences in the folded geometry of foldamers. Such an approach would also give crucial information about foldamer stability and solution dynamics. The necessary condition for having the described predictive ability is to identify minimal foldamer units that contain sufficient relevant structural information. Furthermore, it is critical to have computational tools and parameters that allow for accurate simulation of folded structures in order to study atomistic level mechanisms of functional foldamers in solution.

In the past several years, we have focused on identifying an appropriate computational strategy that allows for accurate prediction of arylamide foldamer structures.[14–18] The essence of this strategy is to take into consideration

specific foldamer building block effects and non-covalent interactions between adjacent monomers that are crucial for arylamide foldamer structure control.

In this study, we present results of an application of this approach to a series of arylamide backbones and demonstrate its structure predictive power. We also present a number of geometrical and dynamical properties obtained from MD simulations of various helical arylamides. The stabilization of the helical structures in the studied series of foldamers is achieved by intramolecular H-bonding between the amide H and an H-bond acceptor on the aromatic entity. We demonstrate here information transferability between the building blocks and the final foldamer and provide the relevant information on these building blocks. Our analysis identifies the minimal structural unit(s) sufficient to reveal the geometrical information for the prediction of the final foldamer structure. These findings will assist the design of foldamers in a more rational and predictable manner.

Table 1. Arylamide elements used in this study and associated re-parameterized torsional parameters.

Name[a]	aryl-N_p	V_n	aryl-C_p	V_n[b]
OMe[14]		8.2 (Y=OCH$_3$)		8.2 (Y=OCH$_3$)
F[15]		6.0 (Y=F)		7.5 (Y=F)
H[c]		3.6 (Y=H)		3.7 (Y=H)
Napy (i)		12.4		6.6
(ii)		8.2		9.6
Py		8.4		8.6
Qn (i)		11.4		7.2
(ii)		8.0		11.2
Bf (i)		5.6		7.8
(ii)		3.2		7.4
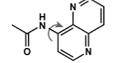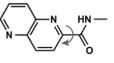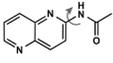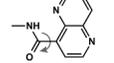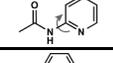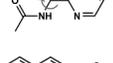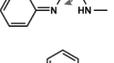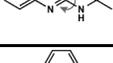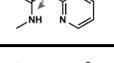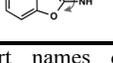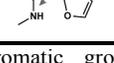
[a] Short names denoting the aromatic groups: **OMe** for methoxybenzene (Y=OCH$_3$), **F** for fluorobenzene (Y=F), **H** for benzene (Y=H), **Py** for pyridine, **Napy** for naphthyridine, **Qn** for quinoline and **Bf** for benzofuran. For each of the double-ring systems (**Napy**, **Qn** and **Bf**), there are four different aryl-amide bond types due to the asymmetry of the two linking positions.

[b] V_n in kcal/mol, $E_{torsion} = \dfrac{V_n}{2}(1+\cos(n\phi - f))$, n = 2 and f = 180°.

[c] V_n for Y=H is given as reference. Note that the parameter given for the aryl-C_p torsion is also re-parameterized, as compared to V_n=29.0 kcal/mol in the original general AMBER force field (GAFF).[17,19]

2. Computational methods

A series of arylamide building blocks were constructed by connecting peptide groups with frequently used aromatic rings, through either aryl-$C_{peptide}$ (C_p) or aryl-$N_{peptide}$ (N_p) bonds, at different positions to provide a systematic set of arylamide foldamer elements (Table 1). The choice allows for comparison with experimental data, available for a number of these aromatic/peptide combinations.

2.1 Torsional parameters and simulation protocols

We have re-parametrized a series of aryl-N_p and aryl-C_p torsional parameters by taking into consideration monomer specific effects such as the influence of intramolecular H-bonding. The re-parameterization process was described in detail in our previous works.[14–16,18] Briefly, we use a high level quantum mechanical (QM) method to calculate the aryl-C_p and aryl-N_p torsional potential energy profiles of model compounds (Table 1). Based on the QM torsional energy profiles and non-bonded interaction energies calculated using the van der Waals parameters in GAFF[19] and RESP charges,[20] we re-parameterize the torsional parameters (Table 1). For a number of model compounds, molecular dynamics (MD) simulations in various solvents were carried out and their conformational behaviour in solutions was validated by nuclear Overhauser effect (NOE) NMR experiments.[15,16,18]

We have also constructed arylamide residues/building blocks (Figure 1 and Figure 2 in brackets) in a way similar to amino acid residues in protein simulations. The partial atomic charges for these residues were derived using the RESP method[20] and restraining the sum of charges of the boxed groups (Figure 1) to zero. AMBER[21] library (.lib) files, containing topology and partial charges, were created for these residues/building blocks. Similarly to building biopolymers, an arylamide foldamer can be built using a simple *sequence* command in LEaP[21] to link a series of residues/building blocks.†
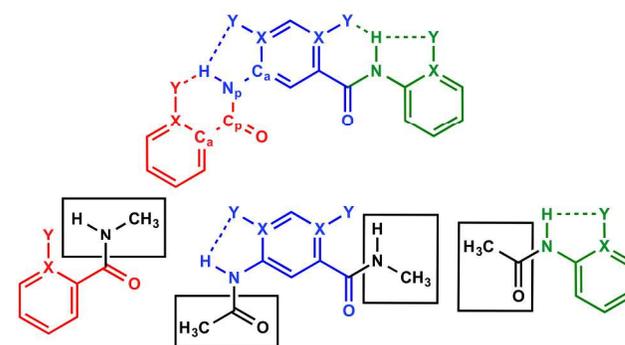
Figure 1. Schematics of arylamide building blocks for head (red), middle (blue) and tail (green) units. The bottom row depicts the molecules with which RESP charges are calculated. The sum of the charges of each boxed group is 0.

2.2 Molecular dynamics simulations

MD simulations were carried out on four types of homo-arylamides built as tetramers, octamers and dodecamers. As shown in Figure 2, type **I** and **II** helices feature conformation-stabilizing intramolecular H-bonds on the outer rims. Type **III** and **IV** helices have the H-bonds on the inner rims. Another classification relates to the difference in the direction and relative orientation of aryl-amide linkages along the backbone. Both type **I** and **III** helices feature the N_p-aryl-C_p units along the entire backbone. This results in a one-directional pattern of the $C_p \rightarrow N_p$ bonds from the N to the C terminus. On the other hand, type **II** and **IV** helices feature both the C_p-aryl-C_p and N_p-aryl-N_p units, which produce alternating $C_p \rightarrow N_p$ and $N_p \rightarrow C_p$ orientations of the peptide bonds along the backbone. Thus, whereas types **I** and **III** require only one kind of building blocks (Figure 2, top, bracketed), two types of building blocks (Figure 2, bottom, blue and pink) were used alternatively in the sequence of type **II** and **IV** helices.

We use the following naming system for each helical foldamer: *oligomeric state - type - aromatic group*. *Oligomeric states* are **Tetra** (tetramer), **Octa** (octamer) and **Dodeca** (dodecamer). *Type* refers to the type of linkage pattern (**I**, **II**, **III**, **IV**, Figure 2) and *aromatic groups* are **OMe**, **F**, **Py**, **Napy**, **Qn** and **Bf** (Table 1). For example, an octamer of quinoline based arylamide of type **III** is named **Octa-III-Qn**. All oligomers are built using the LEaP program by linking a series of building blocks. Idealized left-handed helices (Figure S1) are constructed by tuning the X-C_a-C_p-N_p and X-C_a-N_p-H dihedral angles (Figure 1) of the building blocks to slightly negative (-4° to -8°). Three different explicit solvents (chloroform, methanol and water) have been used in the MD simulations of each oligomer. For tetramer systems, each simulation box contains one oligomer and ~700 (chloroform), ~1500 (methanol) or ~3000 (TIP3P water) solvent molecules and is ~45 Å long on each side. The non-polarizable models for methanol and chloroform provided in the AMBER[21] software package were used. For octamers and dodocamers, the number of solvent molecules roughly doubles and the simulation box measures ~55 Å along each side. Periodic boundary conditions were used with the particle mesh Ewald (PME) procedure for the long-range electrostatic interactions in the periodic box systems. SHAKE algorithm for constraining all bonds involving hydrogen atoms was applied with an integration time step of 1 fs. Each system was first equilibrated for 1-2 ns in an NPT ensemble using anisotropic pressure scaling at constant temperature of 300 K and pressure of 1 atm. Production runs were then performed for 10-20 ns each in an NVT ensemble using the weak-coupling algorithm at constant temperature of 300 K. All simulations were performed using the AMBER 11 software package.[21]

2.3 Conformational analysis outline

Two types of conformational analyses were carried out based on the 10-20 ns MD trajectory (10,000-20,000 conformations) for each system. The first one is at the atomistic level, focusing on the conformation determining backbone dihedral angles, X-C_a-C_p-N_p and X-C_a-N_p-C_p (Figure 1). The conformation of these dihedral angles primarily depends on the characteristics of the intramolecular H-bonding interaction between the *ortho* substituent (OCH_3 or F) or the endocyclic H-bond acceptor ($N_{a(romatic)}$ or $O_{a(romatic)}$) and the amide H, which is the main driving force for the folding of these helices. The distribution of X-C_a-C_p-N_p and X-C_a-N_p-C_p dihedral angles centred at ~0° and ~180°, respectively, corresponds to H-bonded conformations. We therefore use the values of these atomistic dihedral angles, in combination with measurements of the secondary structure properties described below, to assess helix stabilities and separate conformations (Table S1) that maintain helical shape from those that fall apart.
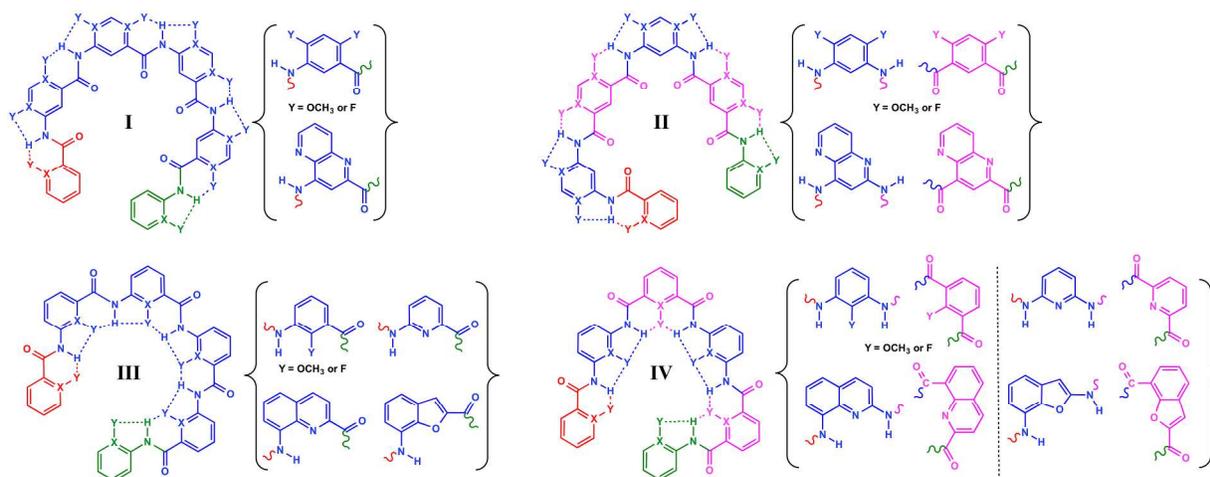
Figure 2. Schematics of the four types of helical arylamide foldamers and residues/building blocks (in brackets) used in each type. Dotted lines denote H-bonds. In the heteroaromatic systems, **Py**, **Napy**, **Qn** and **Bf**, with no exocyclic substituent Y, H-bonds are between the amide H with the endocyclic X (N or O) atom.

The other type of analysis relies on five secondary structure properties obtained from MD simulations. These properties are helix curvature angle, pitch dihedral angle, helical pitch, number of units (residues) per turn (NUpT), and helical pore diameter. We define the helix curvature angle and the pitch dihedral angle by the centres of mass (COMs) of three and four consecutive aromatic rings, respectively. The sign of a pitch dihedral angle indicates the handedness of the helix. Although we don't discuss helix handedness in detail here, it is an important aspect of foldamer research[22,23] and the values of pitch dihedral angle are therefore given for reference.

Determinations of the three remaining secondary structure properties are correlated with each other. We start by defining helical pitch as the distance between the COMs of any two groups that stack on top of each other, where groups are either aromatic rings or peptide bonds. The two stacking groups are one turn apart from each other in the sequence; therefore, by determining the indices of the stacking groups, we also find NUpT. Given that each unit/residue includes one aromatic ring and one peptide group, we assign integer indices $(1, 2, 3, …, n)$ to aromatic rings and "half" numbers $(1.5, 2.5, …, (n-0.5))$ to peptide bonds, starting from the N terminus (n = 4, 8, or 12). Then we calculate the distances between any ith and $(i + j)$th group COMs, in which i = 1, 1.5, ..., n-j. Several j values are used based on the NUpT range, estimated by visual inspection of the MD trajectory. By comparing the peak positions of the distributions of the these COM distances for different j values, we determine the NUpT as the j value that gives the shortest distance between the ith and $(i+j)$th COMs (indicating optimal stacking). Helical pitch is then calculated based on the distribution of distances between COMs of the ith and $(i+NUpT)$th groups. Finally, the helical pore diameter is measured as the distance between the COMs of groups that are half turn (NUpT/2) apart.

For each secondary structure property in an oligomer, multiple distributions are obtained; e.g. there are $(n − 2)$ curvature angles, $(n − 3)$ pitch dihedral angles and $(n − NUpT)$ helical pitches for each oligomer (n = 4, 8 or 12). These distributions are based on helical conformations extracted from MD trajectories. For example, Figure 3 shows the time evolution of the five helical pitches - distances between COMs of the ith and $(i+7)$th aromatic rings - of **Dodeca-I-F** whose NUpT is determined as 7. Stable helical structures are those with all five distances fluctuating around 4 Å (from ~12 ns to 20 ns). As expected, these are also the conformations with the atomistic $X-C_a-C_p-N_p$ and $X-C_a-N_p-C_p$ dihedral angles in the range that would favour the intramolecular H-bonds.

For each oligomer, the distributions of secondary structure properties are largely normal and the multiple curves (Figure 4) of one property overlap with each other. Therefore, we fitted the average distributions of the curvature angles, pitch dihedral angles, helical pitches and pore diameters to Gaussian functions. Tables 2-5 show the secondary structure properties obtained from the Gaussian fit of the average distributions based on all helical conformations (Table S1) extracted from the MD trajectories. They provide quantitative means for comparison of secondary structure features among helical arylamides with varying aromatic groups and linkage types.
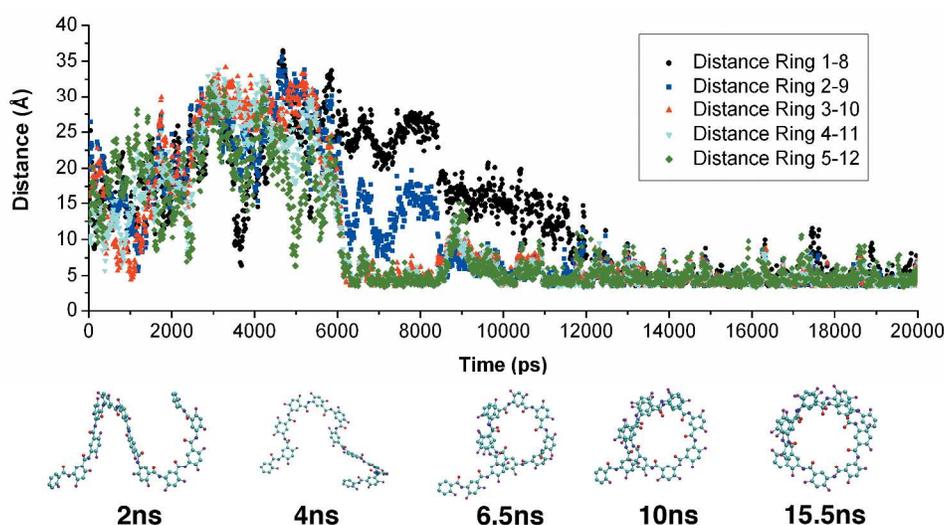
Figure 3. An example of the MD trajectory – distances between COMs of the ith and $(i+NUpT)$th aromatic rings are shown as a function of simulation time for the 20 ns NVT production run for **Dodeca-I-F** in methanol. Snapshots of the dodecamer at selected time intervals are also shown. The left most aromatic ring in the snapshot is ring 1.

3. Results and discussion

The primary goal of this study is conformation prediction by establishing information transferability from building blocks to foldamers. In other words, we identify the minimal structural unit(s) that can be used for quantitative prediction of foldamer secondary structure properties. In sections 3.1 to 3.4, we present the detailed simulation results and analyses, as well as comparison with experimental data where available, for each helical type (Tables 2-5). In section 3.5, we discuss our general findings on information transferability.

3.1 Type I helical arylamide

As shown in Figure 2, to build type **I** helical arylamides, we used three kinds of aromatic groups: *ortho*-OCH$_3$ (**OMe**) and *ortho*-F (**F**) substituted benzene, and naphthyridine (**Napy**). For all three sub-types, one could, based on general chemistry knowledge and intuition, make certain predictions: the curvature angle should be close to but somewhat larger than 120°; the NUpT should be somewhat larger than 6. The considerations that go into the intuitive predictions are that (1) the aryl-N$_p$ and aryl-C$_p$ linkages are *meta* to each other on the six-membered aromatic ring; and (2) the H-bonds on the outer rim of the helix should cause an increase of the curvature angle and consequently the NUpT. Our simulation results (Table 2) verify this intuitive prediction; more importantly, they give quantitative prediction on the significant deviation of curvature angles from 120°. It should be noted that we also distinguish small but significant differences brought by different aromatic groups. For example, secondary structure properties obtained for aqueous solution (Table 2) indicate that **I-OMe** and **I-F** helices have a slightly different curvature angle with NUpT = 7, whereas **I-Napy** has NUpT = 8. These differences cause distinct pore diameters of 13.9, 14.3 and 19.1 Å (Table 2) for **I-OMe**, **I-F** and **I-Napy** helices, respectively.

Comparison of MD simulation results for tetramers, octamers and dodecamers in three solvents also reveals the interplay of the intramolecular H-bonding interactions, solvent effect and aromatic stacking. For tetramers, the strand is not long enough to form one turn and allow for π−π stacking. Thus, analysis of the atomistic dihedral angles X-C$_a$-C$_p$-N$_p$ and X-C$_a$-N$_p$-C$_p$ shows similar behaviour to monomeric compounds.[15,18] Upon dissolution, we observe higher retention rates for H-bonding between the amide H and the endocyclic N$_a$ atoms (>90% for all solvents) than for H-bonding with the exocyclic OCH$_3$ and F substituents. This suggests stronger intramolecular H-bonds in the **I-Napy** oligomers than in **I-OMe** or **I-F**. For **OMe** and **F**, protic solvents decrease the probability of intramolecular H-bonds between the amide proton and the *ortho*-OCH$_3$/*ortho*-F substituents due to competition.[15,18] The percentage of fully H-bonded conformations is >90% in chloroform, whereas it decreases to ~85% and ~57% in water for **Tetra-I-OMe** and **Tetra-I-F**, respectively. The H-bond retention rate is higher for **Tetra-I-OMe** than for **Tetra-I-F** due to the fact that OCH$_3$ is a better H-bond acceptor.

Simulations of **Dodeca-I-OMe** result in stable helical structures in methanol and water ~99.5% of the time. In chloroform, we obtain only partially folded structure with roughly one turn; the rest of the strand has a random conformation. For **Dodeca-I-F**, helical structure is observed 91%, 46% and 88% of the time in chloroform, methanol and water, respectively. The mixed effects by solvent on the stabilities of the **OMe** and **F** dodecamers can be explained by the following rationale. Protic solvents compete with intramolecular H-bond to destabilize the helical structure held by the H-bonds (e.g. 91% to 46% decrease in helical **Dodeca-I-F** from chloroform to methanol). However, they would also induce solvophobic effects, i.e. they cause an increased tendency for aromatic and peptide groups to remain stacked, resulting in intensified π−π stabilization of the helical structure (e.g. 46% to 88% increase in helical **Dodeca-I-F** from methanol to water).

For **Dodeca-I-Napy**, stable helical structures are obtained 100% of the time in all three solvents due to higher retention rates of intramolecular H-bonds. The use of the endocyclic aromatic group also results in helices with NUpT = 8, higher than the 7-7.5 units per turn of the exocyclic **OMe** and **F** systems. The results are consistent with stronger H-bonds in the **I-Napy** oligomers as stronger H-bonds on the outer rim of helix would result in higher curvature angles.

Table 2. Secondary structure properties of type **I** helices.[a]

	Solvents	CA* (°)	Pitch (Å)	NUpT§	Diameter‡ (Å)	PD# (°)
Dodeca-I-OMe	MeOH	129.4 (3.5)	3.9 (0.5)[b]	7-7.5	14.8 (0.7)[c]	-6.7 (17.9)
	H$_2$O	126.6 (3.9)	3.8 (0.3)	7	13.9 (0.8)	-7.1 (14.6)
Dodeca-I-F	CHCl$_3$	129.9 (4.4)	5.0 (1.2)[b]	7-7.5	15.2 (0.9)[c]	-6.5 (23.9)
	MeOH	129.1 (4.1)	3.9 (0.4)	7	14.9 (1.0)	-6.7 (20.8)
	H$_2$O	127.9 (4.0)	3.7 (0.3)	7	14.3 (0.7)	-7.1 (17.1)
Dodeca-I-Napy	CHCl$_3$	133.6 (3.8)	3.9 (0.3)	8	19.4 (0.8)[d]	-5.5 (23.7)
	MeOH	133.6 (3.9)	3.8 (0.3)	8	19.3 (0.8)[d]	-5.8 (24.8)
	H$_2$O	133.3 (3.7)	3.7 (0.2)	8	19.1 (0.8)[d]	-6.2 (22.8)

* Curvature Angle; § Number of Units per Turn; ‡ Pore Diameter; # Pitch Dihedral Angle
[a] Except for NUpT, peak position and standard deviation (in parenthesis) from Gaussian fit are shown.
[b] Obtained from average distribution of distances between COMs of the *i*th and (*i+7*)th groups.
[c] Obtained from average distribution of distances between COMs of the *i*th and (*i+3.5*)th groups.
[d] Obtained from average distribution of distances between COMs of the *i*th and (*i+4*)th aromatic groups.

Interestingly, for both **Dodeca-I-OMe** and **Dodeca-I-F**, we have also observed a clear trend of solvent effect on helical structures. Specifically, the increase of solvent polarity (chloroform → methanol → water) is accompanied by a slight decrease in curvature angle, pitch, NUpT and helical pore diameter, as well as a slight increase in the absolute value of the pitch dihedral angle (Table 2). The trend is also there for **Dodeca-I-Napy**; however, the difference between solvents is smaller. The results indicate a solvophobic effect in which a foldamer adopts compact helical structure in polar solvent.

The behaviour of the octamers can also be explained by solvophobic effects. For **Octa-I-OMe** and **Octa-I-F**, the system has just enough units to form one turn (one $\pi-\pi$ stacking interaction). Therefore, helical conformations are observed only in the aqueous environment (64% of the time for **OMe** and 37% of the time for **F**) where solvophobic effects are stronger. For chloroform and methanol, the conformations observed are extended and random. For **Octa-I-Napy**, which has an insufficient number of units to form a full turn, crescent shaped conformations are found in all three solvents.

Our results are in an excellent agreement with available experimental data. Gong and co-workers have synthesized and characterized various alkoxybenzene based arylamide oligomers that are analogous to type **I-OMe**. The crystallographic data of a nanomer and NMR NOE study of its solution structure (in $CDCl_3$ and DMSO) gave NUpT $\approx$ 7 and the cavity of 10 Å diameter.[24,25] The 10 Å diameter agrees well with our value of 14-15 Å, considering that the two diameters are differently defined. Our diameter is not based on the inner surface of the cavity, but rather measures the distance between the COMs of two aromatic/peptide groups that are half turn apart, where these COMs are about 2-2.5 Å away from the inner surface of the central cavity. This yields a difference of 4-5 Å between the two measurements.

3.2 Type II helical arylamide

Type **II** oligomers are built from the same aromatic groups as those used in type **I**, with H-bond acceptors residing on the outer rim of the helix. The difference is that type **II** uses symmetrical N_p-aryl-N_p and C_p-aryl-C_p units, which alternate in the sequence. This results in two sets of curvature angles (Table 3), one at ~115° (angles centred at the N_p-aryl-N_p units) and the other at ~142° (angles centred at the C_p-aryl-C_p units). As a result, the distributions of pore diameters peak at two slightly different values, indicating more of an oval than circular shaped pore. The difference between the two diameter groups is small and significant overlaps of distributions between the two groups are observed. Therefore, the Gaussian fit is carried out on the overall average of all distributions, and one peak value of pore diameter is given in Table 3. The NUpT values are similar to those of type **I** since the average of the two curvature angles is close to that of type **I**.

Similar to the type **I** tetramers, the distributions of the atomistic dihedral angles X-C_a-C_p-N_p and X-C_a-N_p-C_p of type **II** tetramers are close to monomeric compounds. Polar protic solvents reduce the population of H-bonded conformations. For the **OMe** and **F** oligomers, the percentage of fully H-bonded conformation is similar to type **I** in corresponding solvent. However, due to the asymmetrical linking positions of the double-ring **Napy** system, there are four types of linkages in **II-Napy** as compared to just two types in **I-Napy** (Table 1 and Figure 2). The retention of H-bonds associated with the **Napy**(i,ii)-C_p and **Napy**(i)-N_p linkages is 100% in all solvents for the **II-Napy** systems, similar to that of **I-Napy** which features only the **Napy**(i) linkages. However, the fully H-bonded conformation for **Tetra-II-Napy** decreases to 69% in water due to the more flexible **Napy**(ii)-N_p linkage (Table 1). The main reason for this flexibility is that the **Napy**(ii)-N_p linkage does not facilitate optimal H-bonding interaction. Specifically, an H-bond between N_a and the amide H would form a four-membered H-bond ring (N_a-C_a-N_p-H, Table 1) system which: (1) has longer distance between N_a and amide H; (2) has smaller N_a---H-N_p angle; and (3) is more prone to protic solvent attacks than the five- (**Napy**(i)-N_p, N_a-C_a-C_a-N_p-H, Table 1) or six- (**Napy**(ii)-C_p, N_a-C_a-C_a-C_p-N_p-H, Table 1) membered H-bond ring systems. Therefore, the term H-bonded conformation here simply refers to the atomistic dihedral angle X-C_a-N_p-C_a at around 180° that would facilitate helical structure and a possible H-bond type attraction.

Table 3. Secondary structure properties of type **II** helices.[a]

	Solvents	CA* (°)	Pitch (Å)	NUpT§	Diameter‡ (Å)	PD# (°)
Dodeca-II-OMe	MeOH	114.8 (3.6)/142.0 (3.9)	3.8 (0.4)[b]	7 – 7.5	14.4 (0.9)[c]	-6.4 (22.2)
	H_2O	113.0 (3.8)/140.6 (4.2)	3.8 (0.4)	7	13.9 (0.9)	-7.0 (18.4)
Dodeca-II-F	MeOH	115.3 (4.1)/143.5 (4.3)	4.1 (0.6)[b]	7 – 7.5	14.7 (1.0)[c]	2.0 (24.2)
	H_2O	113.8 (4.0)/142.4 (4.1)	3.9 (0.4)	7	14.3 (0.9)	-7.0 (20.8)
Dodeca-II-Napy	MeOH	115.3 (4.0)/141.1 (4.3)	3.8 (0.3)	7	15.3 (0.9)	-6.6 (23.7)
		117.8 (3.6)/143.8 (4.0)	3.8 (0.3)	8	18.6 (1.4)	-7.4 (36.4)
	H_2O	115.0 (3.8)/140.9 (4.0)	3.7 (0.3)	7	15.2 (0.8)	-6.4 (22.5)
		117.8 (3.5)/143.9 (4.0)	3.7 (0.3)	8	18.6 (1.4)	-8.5 (36.8)

* Curvature Angle; § Number of Units per Turn; ‡ Pore Diameter; # Pitch Dihedral Angle
[a] Except for NUpT, peak position and standard deviation (in parenthesis) from Gaussian fit are shown.
[b] Obtained from the average distribution of distances between COMs of the *i*th and (*i+7*)th groups.
[c] Obtained from the average distribution of distances between COMs of the *i*th and (*i+3.5*)th groups.

Solvophobic effect again plays a vital role in the stability of helices. Stable helical structures for dodecamers are only observed in methanol and water for all three sub-types. The helical structures are in general more compact in water than in methanol (Table 3). In methanol and water, dodecamers remain helical ~100% of the time for both **Napy** and **OMe**, whereas this ratio is only 32% (methanol) and 86% (water) for **F**, similar to type **I**. For **Dodeca-II-F** in methanol, a right-handed helix (Table 3, positive pitch dihedral angle) is obtained, as the oligomer unfolded and folded back during the 20 ns trajectory. This correlates well with the differences in H-bond retention rates for the three aromatic units. For octamers, $\pi-\pi$ stacking between the terminal groups is observed in water only due to stronger solvophobic influence with respect to other solvents.

For **Dodeca-II-Napy**, two types of helical structures are found in both methanol and water, one with NUpT = 7 and the other with NUpT = 8. Analysis of the MD trajectory shows that NUpT = 7 is present for the first 6-7 ns of the simulation, whereas NUpT = 8 dominates the remaining trajectory. It is clear that the $\pi-\pi$ stacking of the aromatic double-rings is the driving force for NUpT being either 7 or 8 rather than 7.5 in which case aromatic-peptide stacking occurs. A longer MD trajectory is needed to determine the relative stability between the two types of helices.

3.3 Type III helical arylamide

The five different type **III** oligomers can be divided into two groups. The first group includes oligomers with aromatic groups **OMe**, **F** and pyridine (**Py**) in which the aryl-N_p and aryl-C_p linkages are *meta* to each other. Each aromatic group has one H-bond acceptor, residing on the inner rim of the helix and being shared by amide H atoms from the two peptide groups linked with the aromatic group. Therefore, it is expected that the curvature angle should be somewhat smaller than 120° and the NUpT should be somewhat smaller than 6. The second group of oligomers is based on quinoline (**Qn**) and benzofuran (**Bf**) building blocks in which the aryl-N_p and aryl-C_p linkages are on different rings of the double-ring aromatic group. In combination with the inner rim H-bonds, the curvature angle is expected to be somewhat smaller than 60° (**Qn**) and 90° (**Bf**) and NUpT somewhat smaller than 3 (**Qn**) and 4 (**Bf**). As discussed below, and shown in Table 4, our simulation results corroborate intuitive prediction and, importantly, determine the deviations of curvature angles from the ad-hoc predictions to be as large as 15° to 25°. The predictive value can be illustrated by the following example. For type **III** helices, we get a series of possible helical pore diameters (e.g. 6.4, 7.6, 8.9 and 10.0 Å) as a function of the aromatic groups used.

For tetramers of the first group (**OMe**, **F** and **Py**), once again the conformational distributions in terms of the X-C_a-C_p-N_p and X-C_a-N_p-C_p dihedral angles are similar to those of the monomeric compounds. The behaviours of **Tetra-III-OMe** and **Tetra-III-F** are similar to those of their corresponding **I** or **II** tetramers in the same solvent. For **Tetra-III-Py**, 100% H-

bonded conformation is found in chloroform whereas the percentage of fully H-bonded conformation decreases to 12% and 6% in methanol and water, respectively, due to the flexible **Py**-N_p linkage. This is similar to the Napy(ii)-N_p linkage in **II-Napy** due to the weak (or non-existing) H-bond interaction at the **Py**-N_p linkage and its vulnerability to protic solvent attacks.

For the **III-OMe** octamers and dodecamers, only **Octa-III-OMe** maintains 100% folded state in methanol and water. As a result of being on the inner rim of the helix and acting as a shared H-bond acceptor between two amide protons, the O-CH_3 bond is now perpendicular to the aromatic plane and therefore parallel to the helical axis. Significant steric repulsion caused by the O-CH_3 groups between two stacking layers is observed. For NUpT = 5, only octamer which has less than 2 turns can maintain the helical structure. This is possible because the steric repulsion is minimized by the first four O-CH_3 groups remaining above the top helical plane, whereas the last four are below the bottom helical plane. For **Dodeca-III-OMe**, we observe a partially folded structure with about two turns, and extended structures beyond two turns as further aromatic stacking causes an increase in steric repulsion by the O-CH_3 groups between two adjacent helical planes.

Solvophobic effects similar to those in types **I** and **II** are also observed here. Both **Octa-III-OMe** and **Dodeca-III-OMe** adopt extended structures in chloroform and either folded or partially folded structures in methanol and water. For **III-F** oligomers, helical structures with NUpT = 5 are obtained in all three solvents for both octamers and dodecamers. However, in methanol, the helical form only accounts for about 39% and 64% of the time for **Octa-III-F** and **Dodeca-III-F**, respectively. This is again a demonstration of a combined solvent effect as we have observed in type **I** oligomers. In chloroform, intramolecular H-bonding holds the helical structure together, whereas in water, strong solvophobic effect is responsible for the high percentage (~99%) of folded structures. The effect of solvents on secondary structure properties (Table 4) is small for the **III-F** helices. This is likely correlated with the fact that, once the helical structure is formed, the substituents are inside the helical pore, thus shielded from solvent access.

Despite the fact that **Tetra-III-Py** experiences significant loss of intramolecular H-bonds in polar protic solvents, we have obtained 100% helical conformation for **Octa-III-Py** and **Dodeca-III-Py** in all three solvents. The only effect of the more flexible **Py**-N_p linkage is a 180° flip of one terminal **Py**-N_p bond of **Dodeca-III-Py** in methanol (14% of the time at the N-terminus, Figure 2 (**III**), red unit) and in water (94% of the time at the C-terminus, Figure 2 (**III**), green unit). However, such 180° flip at the terminus (especially C-terminus) has minimal impact on the aromatic stacking and helical structure. The results indicate a larger effect of aromatic stacking on helical stability in the endocyclic **Py** systems.

Table 4. Secondary structure properties of type **III** helices.[a]

	Solvents	CA* (°)	Pitch (Å)	NUpT§	Diameter‡ (Å)	PD$^{#}$ (°)
Octa-III-OMe	MeOH	106.2 (4.7)	3.9 (0.4)	5	10.4 (0.7)	-13.1 (18.8)
	H$_2$O	104.3 (4.6)	4.4 (0.5)	5	9.9 (0.6)	-12.5 (9.0)
Dodeca-III-F	CHCl$_3$	104.8 (3.4)	3.8 (0.4)	5	10.1 (0.4)	-12.9 (12.2)
	MeOH	104.8 (3.3)	3.8 (0.3)	5	10.1 (0.4)	-12.7 (11.7)
	H$_2$O	104.1 (3.0)	3.8 (0.3)	5	10.0 (0.4)	-12.4 (10.5)
Octa-III-F	CHCl$_3$	105.4 (3.8)	3.9 (0.4)	5	10.2 (0.5)	-14.1 (17.5)
	MeOH	105.5 (3.9)	3.8 (0.4)	5	10.2 (0.5)	-13.0 (15.7)
	H$_2$O	104.0 (3.4)	3.9 (0.4)	5	10.0 (0.4)	-12.8 (12.8)
Dodeca-III-Py	CHCl$_3$	99.5 (2.9)	3.5 (0.3)	4.5	9.1 (0.4)	-14.3 (9.2)
	MeOH	99.0 (3.1)	3.5 (0.3)	4.5	9.0 (0.4)	-14.6 (9.0)
	H$_2$O	98.0 (3.1)	3.6 (0.3)	4.5	8.9 (0.4)	-15.0 (7.6)
Octa-III-Py	CHCl$_3$	99.0 (3.1)	3.5 (0.3)	4.5	9.1 (0.4)	-14.4 (13.1)
	MeOH	98.8 (3.2)	3.5 (0.3)	4.5	9.1 (0.4)	-14.3 (13.4)
	H$_2$O	97.6 (3.4)	3.6 (0.3)	4.5	8.9 (0.4)	-14.7 (10.6)
Dodeca-III-Qn	CHCl$_3$	42.3 (3.1)	3.5 (0.2)	2.5	6.4 (0.3)	-61.7 (5.9)
	MeOH	42.2 (3.0)	3.5 (0.2)	2.5	6.4 (0.3)	-61.7 (5.8)
	H$_2$O	41.9 (3.0)	3.5 (0.2)	2.5	6.4 (0.3)	-61.7 (5.8)
Octa-III-Qn	CHCl$_3$	42.6 (3.2)	3.6 (0.2)	2.5	6.4 (0.3)	-61.2 (6.2)
	MeOH	42.4 (3.2)	3.6 (0.2)	2.5	6.4 (0.3)	-61.7 (6.1)
	H$_2$O	42.0 (3.1)	3.5 (0.2)	2.5	6.4 (0.3)	-61.6 (6.0)
Tetra-III-Qn	CHCl$_3$	43.3 (3.4)	3.6 (0.3)	2.5	6.5 (0.3)	-56.4 (7.4)
	MeOH	42.8 (3.4)	3.6 (0.3)	2.5	6.5 (0.3)	-57.8 (7.3)
	H$_2$O	41.6 (3.3)	3.6 (0.2)	2.5	6.4 (0.3)	-59.1 (7.0)
Dodeca-III-Bf	CHCl$_3$	65.1 (5.3)	3.6 (0.2)	3	7.8 (0.3)	-29.0 (4.3)
	MeOH	65.1 (5.3)	3.6 (0.2)	3	7.6 (0.4)	-31.7 (5.2)
	H$_2$O	65.4 (5.5)	3.6 (0.2)	3	7.6 (0.4)	-31.2 (5.4)
Octa-III-Bf	CHCl$_3$	67.2 (4.6)	3.6 (0.2)	3	7.7 (0.3)	-29.3 (4.5)
	MeOH	65.4 (5.8)	3.6 (0.2)	3	7.6 (0.4)	-31.8 (5.8)
	H$_2$O	64.0 (5.5)	3.6 (0.2)	3	7.5 (0.4)	-32.6 (5.5)
Tetra-III-Bf	CHCl$_3$	67.2 (6.9)	3.6 (0.3)	3	7.7 (0.4)	-35.0 (5.5)
	MeOH	68.9 (7.7)	3.7 (0.3)	3	7.7 (0.4)	-36.4 (6.1); 31.1 (6.3)
	H$_2$O	62.7 (5.9)	3.6 (0.3)	3	7.4 (0.4)	-36.2 (4.9); 31.1 (4.7)

* Curvature Angle; § Number of Units per Turn; ‡ Pore Diameter; # Pitch Dihedral Angle

[a] Except for NUpT, peak position and standard deviation (in parenthesis) from Gaussian fit are shown.

One significant difference between the **III-Py** and **III-OMe**, **III-F** systems is that NUpT of **III-Py** is 4.5, whereas NUpT of **III-OMe**, **III-F** is 5. Likewise, the curvature angle, pitch and pore diameter are smaller than those of the corresponding **OMe** and **F** oligomers. Stronger aromatic stacking and intramolecular H-bonds (note the exception at the Py-N$_p$ and Napy(ii)-N$_p$ linkages) in the heteroaromatic systems should both contribute to the more compact helix. As stronger H-bonds on the outer rim of the helix lead to larger NUpT (for type **I** or **II Napy** systems), stronger H-bonds on the inner rim of the helix result in smaller curvature angles and more compact helical structure (for **III-Py**).

The helical type **III-Qn** systems have been studied extensively by Huc et al.[11,26] due to: (1) well established synthesis routine for oligomers of any length; (2) well characterized stable helical conformations in solid state and solution (both polar and nonpolar solvents); and (3) helical stability being very sensitive to oligomer length and temperature, which makes this system very useful for studying both dynamic and kinetic properties of the synthetic foldamers. Stable helical structures are obtained for all three lengths of oligomers and all three solvents 100% of the time. The secondary structure properties (Table 4) obtained from our simulations agree very well with experimental measurements by single-crystal X-ray diffraction of an octamer in terms of pitch (~3.5 Å) and number of units per turn (2.5).[11,26] Similarly to the other systems discussed above, these properties are only slightly affected by solvent types.

The **III-Bf** oligomers are based on a 6+5 membered aromatic rings, which results in a unique curvature angle of ~65°. Consequently, the NUpT and pore diameter increase to 3 and ~7.7 Å from the 2.5 and ~6.5 Å of the 6+6 quinoline systems, respectively. It is clear that the **III-Bf** helices are less stable than the **III-Qn** ones due to the more flexible **Bf**-amide

linkage (Table 1). Whereas for most **III-Bf** systems, >90% conformations remain helical (Table S1), helical structure is obtained only 56% of the time for **Tetra-III-Bf** in methanol. In addition, a ~ 40/60 mixture of left-handed and right-handed conformations is obtained for **Tetra-III-Bf** in both methanol and water. Handedness conversion occurs on the scale of <1 ns in methanol. This scale ranges from 2 to 8 ns in water.

3.4 Type IV helical arylamide

The results on type **IV-OMe** oligomers are very similar to those of **III-OMe** ones. Helical structures are only obtained for **Octa-IV-OMe** in water where there are less than two turns in the helix. For all other aromatic groups (**F, Py, Qn** and **Bf**), stable helical structures are obtained when $n > $ NUpT. Similarly to type **II**, there are two sets of curvature angles with the average of the two close to that of type **III**. This time, the curvature angle centred at the N_p-aryl-N_p linkage is larger than the one centred at the C_p-aryl-C_p linkage. The resulting oval shape of the helix pore is more pronounced because of the smaller curvature angles than in type **II**. Therefore, the distributions of pore diameters are divided into two sub groups

and each group is fitted separately (Table 5). For example, for **Octa-IV-F** with NUpT = 5, the diameter of ~10.6 Å corresponds to distances from the COMs of aromatic groups with the aryl-C_p linkages to their corresponding +2.5 or -2.5 peptide groups. Li et al. have synthesized a fluorobenzene based heptamer, analogous to **IV-F**, and characterized its solution structure by 2D NOE experiments.[27] Their data strongly support a helical conformation in chloroform. Our result is in line with their reported cavity diameter of ~6.5 Å of the heptamer, considering the difference in the measurement method, discussed in section 3.1. The smaller diameter of ~9.6 Å relates to distances from the COMs of aromatic groups with the aryl-N_p linkages to their corresponding +2.5 or -2.5 peptide groups. The endocyclic type **IV-Py** oligomers again have more compact helical structures (NUpT = 4.5) than the exocyclic type **IV-F** systems. Notably, the NUpT and pitch obtained for the **IV-Py** oligomers are in excellent agreement with experimental measurements by Huc et al.[28,29] Similarly to **III-Py**, for the **IV-Py** helices in methanol and water, we have observed 180° flip of the **Py**-N_p bonds at the terminus.

Table 5. Secondary structure properties of type **IV** helices.[a]

	Solvents	CA* (°)	Pitch (Å)	NUpT§	Diameter‡ (Å)	PD# (°)
Dodeca-IV-F	CHCl₃	119.6 (3.5)/91.4 (3.1)	3.8 (0.4)	5	10.6 (0.4)/9.6 (0.4)	-12.6 (12.4)
	MeOH	119.3 (3.4)/91.1 (3.1)	3.8 (0.4)	5	10.6 (0.4)/9.6 (0.4)	-12.8 (11.7)
	H₂O	118.7 (3.2)/90.6 (3.0)	3.9 (0.4)	5	10.5 (0.4) 9.5 (0.4)	-12.5 (10.3)
Octa-IV-F	CHCl₃	119.4 (3.5)/91.5 (3.2)	3.8 (0.4)	5	10.6 (0.4)/9.6 (0.5)	-12.6 (15.4)
	MeOH	119.6 (3.7)/91.6 (3.5)	3.9 (0.4)	5	10.6 (0.4)/9.6 (0.5)	-13.2 (19.4)
	H₂O	118.4 (3.5)/90.5 (3.3)	3.9 (0.4)	5	10.5 (0.4)/9.5 (0.4)	-13.1 (13.0)
Dodeca-IV-Py	CHCl₃	117.9 (3.1)/84.9 (3.7)	3.6 (0.3)	4.5	9.9 (0.4)/8.7 (0.5)	-14.2 (11.7)
	MeOH	117.6 (3.6)/83.2 (3.5)	3.6 (0.3)	4.5	9.9 (0.4)/8.5 (0.5)	-14.3 (13.9)
	H₂O	115.8 (3.4)/80.8 (2.6)	3.6 (0.3)	4.5	9.6 (0.4)/8.2 (0.4)	-15.1 (10.5)
Octa-IV-Py	CHCl₃	117.9 (3.4)/84.4 (3.7)	3.6 (0.3)	4.5	9.9 (0.4)/8.6 (0.5)	-14.1 (16.2)
	MeOH	117.5 (3.9)/83.2 (3.4)	3.6 (0.3)	4.5	9.9 (0.4)/8.5 (0.5)	-14.3 (16.1)
	H₂O	115.6 (3.5)/80.6 (2.7)	3.6 (0.3)	4.5	9.5 (0.4)/8.1 (0.4)	-15.1 (11.7)
Dodeca-IV-Qn	CHCl₃	59.1 (3.8)/37.4 (3.2)	3.4 (0.2)	2.5	7.5 (0.3)/6.0 (0.3)	-53.6 (6.4)
	MeOH	59.4 (4.1)/37.8 (3.4)	3.4 (0.2)	2.5	7.6 (0.3)/6.0 (0.3)	-53.0 (6.6)
	H₂O	60.0 (4.4)/37.4 (4.1)	3.4 (0.2)	2.5	7.6 (0.3)/6.0 (0.3)	-51.5 (6.8)
Octa-IV-Qn	CHCl₃	58.9 (4.0)/37.3 (3.5)	3.5 (0.2)	2.5	7.5 (0.3)/6.0 (0.3)	-53.6 (6.6)
	MeOH	59.1 (4.4)/37.6 (3.7)	3.5 (0.2)	2.5	7.5 (0.3)/6.0 (0.3)	-53.3 (7.0)
	H₂O	59.1 (4.9)/37.4 (4.1)	3.5 (0.2)	2.5	7.6 (0.3)/6.0 (0.3)	-52.4 (7.2)
Tetra-IV-Qn	CHCl₃	56.6 (4.0)/35.0 (3.1)	3.5 (0.2)	2.5	7.4 (0.4)/5.9 (0.3)	-58.2 (7.4)
	MeOH	55.4 (3.6)/34.6 (3.0)	3.6 (0.2)	2.5	7.3 (0.4)/5.8 (0.3)	-59.8 (7.0)
	H₂O	54.0 (3.2)/33.8 (2.7)	3.6 (0.2)	2.5	7.2 (0.4)/5.8 (0.3)	-60.7 (6.5)
Dodeca-IV-Bf	CHCl₃	83.9 (3.1)/54.0 (2.7)	3.8 (0.4)	3	8.6 (0.2)/6.8 (0.2)	-29.0 (3.4)
	MeOH	83.2 (3.3)/53.5 (2.8)	3.8 (0.4)	3	8.5 (0.2)/6.7 (0.2)	-29.8 (3.4)
	H₂O	82.5 (3.3)/53.2 (2.8)	3.8 (0.4)	3	8.5 (0.2)/6.7 (0.2)	-30.0 (3.4)
Octa-IV-Bf	CHCl₃	84.1 (3.2)/53.9 (2.8)	3.9 (0.4)	3	8.5 (0.2)/6.8 (0.3)	-29.1 (3.6)
	MeOH	83.7 (3.3)/53.6 (2.9)	3.9 (0.4)	3	8.5 (0.2)/6.7 (0.3)	-29.8 (3.8)

		82.9 (3.4)/53.1 (3.0)	3.8 (0.4)	3	8.5 (0.2)/6.7 (0.2)	-30.0 (3.6)
	H$_2$O	82.9 (3.4)/53.1 (3.0)	3.8 (0.4)	3	8.5 (0.2)/6.7 (0.2)	-30.0 (3.6)
Tetra-IV-Bf	CHCl$_3$	85.3 (4.4)/53.7 (3.8)	4.3 (0.5)	3	8.6 (0.3)/6.8 (0.3)	-33.5 (6.4)/27.4 (6.1)
	MeOH	85.9 (5.0)/53.4 (4.4)	4.3 (0.5)	3	8.6 (0.3)/6.7 (0.3)	-34.7 (6.7)/28.5 (6.1)
	H$_2$O	82.4 (5.1)/51.6 (3.5)	4.0 (0.4)	3	8.4 (0.3)/6.6 (0.3)	-33.0 (5.3)/27.4 (5.3)

* Curvature Angle; § Number of Units per Turn; ‡ Pore Diameter; # Pitch Dihedral Angle
[a] Except for NUpT, peak position and standard deviation (in parenthesis) from Gaussian fit are shown.

Apart from the two sets of curvature angles and helix pore diameters, the results on type **IV** oligomers with the double-ring **Qn** and **Bf** aromatic units are similar to those of type **III**. For type **IV-Qn**, the helical conformation is maintained 100% of the time despite the presence of the more flexible **Qn**(ii)-N$_p$ linkage. This is reasonable considering that **IV-Qn** has the most compact (NUpT = 2.5) helical structure, therefore it is more stabilized by the $\pi-\pi$ interactions, as compared to **III-Py** and **II-Napy** with similar aryl-N$_p$ linkages. For type **IV-Bf** oligomers, the **Bf**(ii)-N$_p$ linkage is the most flexible in terms of both H-bond strength (reflected by the torsional parameter in Table 1) and its vulnerability to solvent attacks. The compensation from its strong $\pi-\pi$ interactions results in the following. For **Tetra-IV-Bf**, helical conformation accounts for 86%, 21% and 40% of the time during the 20 ns trajectory in chloroform, methanol and water, respectively. Handedness inversion occurs with a frequency ranging from <1 ns to ~7 ns. For **Octa-IV-Bf** and **Dodeca-IV-Bf**, helical conformations are maintained ~100% of time in all solvents. However, the 180° flip of the **Bf**(ii)-N$_p$ linkage is observed for about 22%/55% (**Octa/Dodeca**) of the time in methanol and 6% (**Dodeca**) of the time in water.
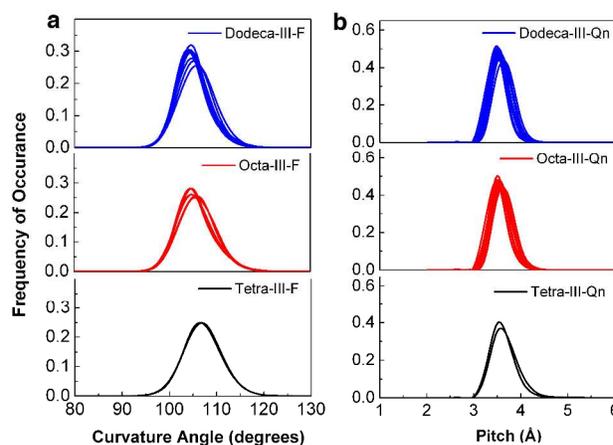
Figure 4. Examples of distributions of secondary structure properties. (a) Distributions of curvature angles **III-F** oligomers in chloroform. (b) Distributions of helical pitches of **III-Qn** oligomers in chloroform.

3.5 Information transferability

Our results clearly demonstrate that the curvature, which varies largely with the type of building blocks, can be determined by MD simulations of a unit as small as a tetramer, i.e. even with less than one helical turn. As shown in Table 6 and Figure 4a, the distributions of curvature angles from MD simulations of tetramer, octamer and dodecamer coincide, indicating that the curvature angle determined at the tetramer

ARTICLE

level can be used to predict the curvature of larger helices. For the **III-F** oligomers shown in Figure 4a, the tetramer has less than one turn; yet it accurately captures the curvature angle of a helix with several turns. This is a consequence of the fact that curvature angles are mainly affected by local structural features such as linkage types and positions, and intramolecular H-bonds. We obtained a large variety of curvature angles, ranging from ~33° to ~133°, using the aromatic/peptide combinations in Table 1.
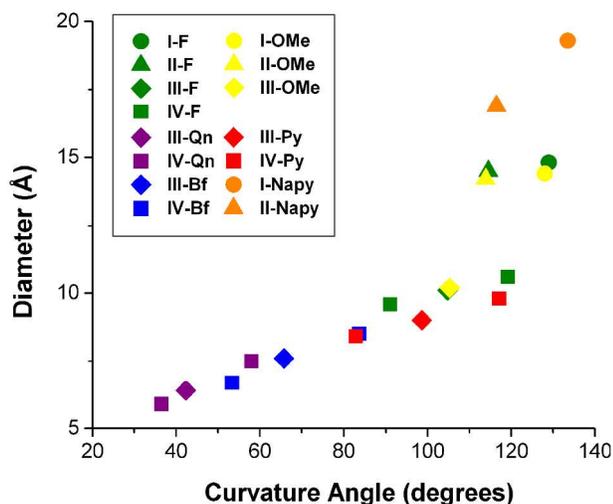
Figure 5. Correlation between curvature angles and pore diameters. Each plotted curvature angle/pore diameter value for each oligomer type is an overall average among all oligomeric states with n > NUpT and three solvents.

The minimal structural units that lead to accurate prediction of pitch related properties are those with at least one turn (Figure 4b and Tables 2-5). Helical pitch is mainly determined by the π−π interactions between the two stacking aromatic or peptide groups. Thus, we observe small variance in pitch among all helical oligomers (mostly from 3.5 to 3.9 Å). In general, aromatic building blocks with exocyclic OCH_3 or F substituents tend to have slightly larger pitch (~3.8 to 3.9 Å) than the heteroaromatic systems (~3.5 to 3.6 Å). Pitch dihedral angle correlates with both pitch and the curvature angle. Given the small variance in pitch, the obtained large variance of the pitch dihedral angle (from ~ -5° to ~ -61°) is a direct result of the large variance in the curvature angles.

The small variance in helical pitch among different arylamide helices leads to the outcome that diversity in helical structures is related mostly to the large variance in curvature angle. As shown in Figure 5, curvature angle can be controlled using different building blocks (e.g. single vs. double aromatic ring, six- vs. five-membered aromatic ring) and linkage positions (*meta-* vs. *para-*, positioned at a single ring vs. across double-rings). Figure 5 also demonstrates the correlation between pore diameter and the curvature angle; the same is the case with other secondary structure properties. Larger curvature angle leads to smaller absolute value of pitch dihedral angle, larger NUpT and larger pore diameters. Therefore, curvature angle, which can be predicted quantitatively based on relatively simple MD simulations of tetramers, turns out to be the crucial secondary structure feature in an arylamide foldamer.

Table 6. Comparison of the curvature angles (°) at different oligomeric states.[a]

	Solvents	Tetramer	Octamer	Dodecamer
I-OMe	CHCl₃	129.4 (4.3)	129.3 (4.6)	X[b]
	MeOH	129.6 (4.2)	129.5 (4.2)	129.4 (3.5)
	H₂O	129.5 (4.2)	123.1 (7.1)	126.6 (3.9)
I-F	CHCl₃	130.0 (4.5)	129.9 (4.5)	129.9 (4.4)
	MeOH	129.7 (4.6)	129.8 (4.7)	129.1 (4.1)
	H₂O	129.4 (4.9)	125.3 (6.2)	127.9 (4.0)
I-Napy	CHCl₃	135.2 (4.3)	135.0 (4.3)	133.6 (3.8)
	MeOH	134.6 (4.5)	135.0 (4.2)	133.6 (3.9)
	H₂O	135.0 (4.3)	134.2 (4.7)	133.3 (3.7)
II-OMe	CHCl₃	116.0 (4.0)/142.7 (4.6)	116.5 (4.1)/142.1 (5.0)	X
	MeOH	116.3 (3.9)/142.4 (4.6)	116.3 (4.0)/142.6 (4.6)	114.8 (3.6)/142.0 (3.9)
	H₂O	116.5 (3.8)/142.6 (4.6)	110.2 (6.3)/136.2 (8.0)	113.0 (3.8)/140.6 (4.2)
II-F	CHCl₃	116.2 (4.4)/144.2 (4.4)	116.2 (4.4)/ 144.2 (4.4)	X
	MeOH	115.9 (4.6)/143.5 (4.9)	116.2 (4.5)/143.5 (4.9)	115.3 (4.1)/143.5 (4.3)
	H₂O	115.7 (4.7)/143.0 (5.1)	111.9 (5.8)/139.7 (6.5)	113.8 (4.0)/142.4 (4.1)
II-Napy	CHCl₃	117.3 (4.5)/143.2 (4.8)	116.9 (4.7)/142.6 (5.1)	X
	MeOH	117.1 (5.0)/142.2 (5.5)	116.6 (4.8)/142.2 (5.3)	115.3 (4.0)/141.1 (4.3)[c] 117.8 (3.6)/143.8 (4.0)[c]
	H₂O	117.2 (4.8)/142.8 (5.1)	109.5 (5.6)/136.6 (6.8)	115.0 (3.8)/140.9 (4.0)[c] 117.8 (3.5)/143.9 (4.0)[c]
III-OMe	CHCl₃	106.9 (4.0)	X	X
	MeOH	107.4 (4.0)	106.2 (4.7)	X
	H₂O	106.9 (4.0)	104.3 (4.6)	X
III-F	CHCl₃	106.9 (4.0)	105.4 (3.8)	104.8 (3.4)
	MeOH	107.4 (4.4)	105.5 (3.9)	104.8 (3.3)
	H₂O	107.8 (4.6)	104.0 (3.4)	104.1 (3.0)
III-Py	CHCl₃	99.1 (3.7)	99.0 (3.1)	99.5 (2.9)
	MeOH	98.2 (4.1)	98.8 (3.2)	99.0 (3.1)
	H₂O	99.0 (4.3)	97.6 (3.4)	98.0 (3.1)
IV-F	CHCl₃	120.6 (4.2)/94.1 (4.1)	119.4 (3.5)/91.5 (3.2)	119.6 (3.5)/91.4 (3.1)
	MeOH	121.1 (4.6)/94.2 (4.6)	119.6 (3.7)/91.6 (3.5)	119.3 (3.4)/91.1 (3.1)
	H₂O	121.4 (4.9)/94.8 (4.6)	118.4 (3.5)/90.5 (3.3)	118.7 (3.2)/90.6 (3.0)
IV-Py	CHCl₃	117.0 (4.0)/83.5 (3.7)	117.9 (3.4)/84.4 (3.7)	117.9 (3.1)/84.9 (3.7)
	MeOH	117.6 (4.7)/83.4 (3.8)	117.5 (3.9)/83.2 (3.4)	117.6 (3.6)/83.2 (3.5)
	H₂O	118.1 (4.4)/82.7 (3.9)	115.6 (3.5)/80.6 (2.7)	115.8 (3.4)/80.8 (2.6)

[a] Peak position and standard deviation (in parenthesis) from Gaussian fit are shown. Similarly to Tables 2-5 where only helical conformations are included in the fitting, for oligomers with n < NUpT, only fully H-bonded conformations are included.
[b] X indicates either no stable helical conformation or no fully H-bonded conformation (for n < NUpT).
[c] Two sets of curvature angles for **Dodeca-II-Napy** in MeOH and H₂O correspond to conformations with 7 and 8 units per turn, respectively. See Table 3 for details.

General chemical structure knowledge can lead us to a rough estimate of curvature angle. For example, in type **III** oligomer of the single-ring systems, curvature angle should be somewhat smaller than 120° due to the intramolecular H-bonds along the inner rim of the helix. This estimate, however, can not be used for deriving an accurate prediction of overall helical parameters, such as diameter or NUpT. MD simulations, on the other hand, can predict curvature angle differences as small as 5°, which in case of type **III** helices results in a difference of 5 vs 4.5 in NUpT and a difference of 1 Å in helix pore diameters (**III-F** vs. **III-Py**).

Our results also indicate that a curvature angle is mainly determined by the characteristics of the middle aromatic group and its two aryl-amide bonds. For example, a curvature angle centered at C_p-**Py**-C_p differs from the one centered at N_p-**Py**-C_p or C_p-**Qn**-C_p (Table 4). In other words, the curvature angle is found to be a highly locally controlled property of arylamide oligomers. This in turn extends the predictive power of our approach beyond the homo-oligomers presented here. Tables 2-5 provide reliable structural predictions for foldamers that can be built from a mixture of building blocks and linkage types. Our study on helical foldamer capsules[30] based on a hetero-

sequence of **Py** and **Qn** building blocks provides validation of the above conclusion.

MD simulations also provide a good description of solvent effects. The results are complex and differ from system to system due to interplay of influence of protic solvents on intramolecular H-bonds, solvophobic effects and $\pi-\pi$ stacking. Although experimental data is not availalble for direct comparison, consistency of the results throughout the different types of oligomers provides us with confidence in the accuracy of the approach.

4. Conclusions

We have utilized a computational approach for the prediction of arylamide foldamer secondary structures in solution. There are several aspects of the predictive power of this approach. The first is that conformational characteristics and behaviour of smaller oligomers closely relate to those of foldamers with longer sequences. We identified tetramer as the minimal unit predictive of curvature angle, and oligomer with at least one turn as predictive of helical pitch. In addition, diameter, pitch dihedral angle and NUpT are predictable from values of curvature angle and pitch. This provides us with the ability to predict geometrical parameters of larger oligomers based on data obtained from simple MD simulations of short oligomers, thus avoiding complex and time consuming simulations of large foldamers.

The second aspect of the predictive power is that conformational information obtained from the simulations of homo-oligomers can give, without running further MD simulations, quantitative structure prediction of arylamide foldamers with mixed building block sequences.

Finally, MD simulations of an actual oligomer provide an accurate description of its conformational behaviour in solution, provided that improved force field parameters are used.

MD simulations on biopolymers, specifically DNA, with systematically varying lengths and monomer compositions have revealed crucial information on flexible vs rigid base pair sequences, as well as important geometrical features that contribute to the shape and functionality of DNA.[31] Taking into account the richer diversity of synthetic foldamers and their growing applications, it is important to have similar understanding of foldamer conformations. Unlike biopolymers, the force field parameters for synthetic oligomers are less robust,[17,18] which hinders such systematic study. This work, based on the improved force field parameters, to the best of our knowledge, is the first systematic computational study of foldamers with various sequence lengths and compositions. Further uses of the computational approach presented here will provide useful insight into dynamical, mechanistic and functional properties of the arylamide oligomer class,[30] which will facilitate rational design of foldamers.

Notes and references

[a] Department of Chemistry & Biochemistry, University of the Sciences, 600 South 43rd Street, Philadelphia, PA 19104, USA.
Email: v.pophri@usciences.edu

† Parameters in AMBER frcmod format and AMBER library files for arylamide residues are available upon request. Electronic Supplementary Information (ESI) available: Percentage of helical conformation/fully H-bonded conformation; Selected initial structures and snapshots from simulations; Examples of MD trajectories with snapshots; See DOI: 10.1039/b000000x/

1. D.-W. Zhang, X. Zhao, J.-L. Hou, and Z.-T. Li, *Chem. Rev.*, 2012, **112**, 5271–316.
2. S. Hecht and I. Huc, Eds., *Foldamers: Structure, Properties and Applications*, John Wiley & Sons, 2007.
3. G. Guichard and I. Huc, *Chem. Commun.*, 2011, **47**, 5933–41.
4. D. M. Chenoweth, D. A. Harki, J. W. Phillips, C. Dose, and P. B. Dervan, *J. Am. Chem. Soc.*, 2009, **131**, 7182–8.
5. J. A. Raskatov, J. L. Meier, J. W. Puckett, F. Yang, P. Ramakrishnan, and P. B. Dervan, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **109**, 1023–1028.
6. J. Garric, J.-M. Léger, and I. Huc, *Angew. Chem. Int. Ed.*, 2005, **44**, 1954–8.
7. J. Garric, J.-M. Léger, and I. Huc, *Chem. - Eur. J.*, 2007, **13**, 8454–62.
8. R. A. Smaldone and J. S. Moore, *Chem. - Eur. J.*, 2008, **14**, 2650–7.
9. W. Cai, G.-T. Wang, Y.-X. Xu, X.-K. Jiang, and Z.-T. Li, *J. Am. Chem. Soc.*, 2008, **130**, 6936–7.
10. Z.-Q. Wu, X.-B. Shao, C. Li, J.-L. Hou, K. Wang, X.-K. Jiang, and Z.-T. Li, *J. Am. Chem. Soc.*, 2005, **127**, 17460–8.
11. N. Delsuc, T. Kawanami, J. Lefeuvre, A. Shundo, H. Ihara, M. Takafuji, and I. Huc, *ChemPhysChem*, 2008, **9**, 1882–90.
12. I. Huc, *Eur. J. Org. Chem.*, 2004, **2004**, 17–29.
13. S. H. Gellman, *Acc. Chem. Res.*, 1998, **31**, 173–180.
14. Z. Liu, A. Teslja, and V. Pophristic, *J. Comput. Chem.*, 2011, **32**, 1846–58.
15. J. F. Galan, J. Brown, J. L. Wildin, Z. Liu, D. Liu, G. Moyna, and V. Pophristic, *J. Phys. Chem. B*, 2009, **113**, 12809–15.
16. J. F. Galan, C. N. Tang, S. Chakrabarty, Z. Liu, G. Moyna, and V. Pophristic, *Phys. Chem. Chem. Phys.*, 2013, **15**, 11883–92.
17. V. Pophristic, S. Vemparala, I. Ivanov, Z. Liu, M. L. Klein, and W. F. DeGrado, *J. Phys. Chem. B*, 2006, **110**, 3517–26.
18. Z. Liu, R. C. Remsing, D. Liu, G. Moyna, and V. Pophristic, *J. Phys. Chem. B*, 2009, **113**, 7041–4.
19. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–74.
20. J. Wang, P. Cieplak, and P. A. Kollman, *J. Comput. Chem.*, 2000, **21**, 1049–1074.
21. D. A. Case, T. A. Darden, T. E. Cheatham, III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Götz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P. A. Kollman, 2011, AMBER 11, University of California, San Francisco.
22. C. Dolain, H. Jiang, J.-M. Léger, P. Guionneau, and I. Huc, *J. Am. Chem. Soc.*, 2005, **127**, 12943–51.
23. A. M. Kendhale, L. Poniman, Z. Dong, K. Laxmi-Reddy, B. Kauffmann, Y. Ferrand, and I. Huc, *J. Org. Chem.*, 2011, **76**, 195–200.
24. B. Gong, H. Zeng, J. Zhu, L. Yuan, Y. Han, S. Cheng, M. Furukawa, R. D. Parra, A. Y. Kovalevsky, J. L. Mills, E. Skrzypczak-Jankun, S.

Martinovic, R. D. Smith, C. Zheng, T. Szyperski, X. C. Zeng, and L. Yua, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 11583–8.

25. B. Gong, *Acc. Chem. Res.*, 2008, **41**, 1376–86.

26. H. Jiang, J.-M. Léger, and I. Huc, *J. Am. Chem. Soc.*, 2003, **125**, 3448–9.

27. C. Li, S.-F. Ren, J.-L. Hou, H.-P. Yi, S.-Z. Zhu, X.-K. Jiang, and Z.-T. Li, *Angew. Chemie*, 2005, **117**, 5871–5875.

28. V. Berl, I. Huc, R. G. Khoury, M. J. Krische, and J. M. Lehn, *Nature*, 2000, **407**, 720–3.

29. V. Berl, I. Huc, R. G. Khoury, and J. M. Lehn, *Chem. - Eur. J.*, 2001, **7**, 2810–20.

30. A. M. Abramyan, Z. Liu, and V. Pophristic, *Phys. Chem. Chem. Phys.*, 2014, **16**, 20406–10.

31. R. Lavery, K. Zakrzewska, D. Beveridge, T. C. Bishop, D. A. Case, T. C. Iii, S. Dixit, B. Jayaram, F. Lankas, C. Laughton, J. H. Maddocks, A. Michon, and R. Osman, M. Orozco, A. Perez, T. Singh, N. Spackova and J. Sponer, *Nucleic Acids Res.,* 2010, **38**, 299–313.

7 units per turn   **5 units per turn**
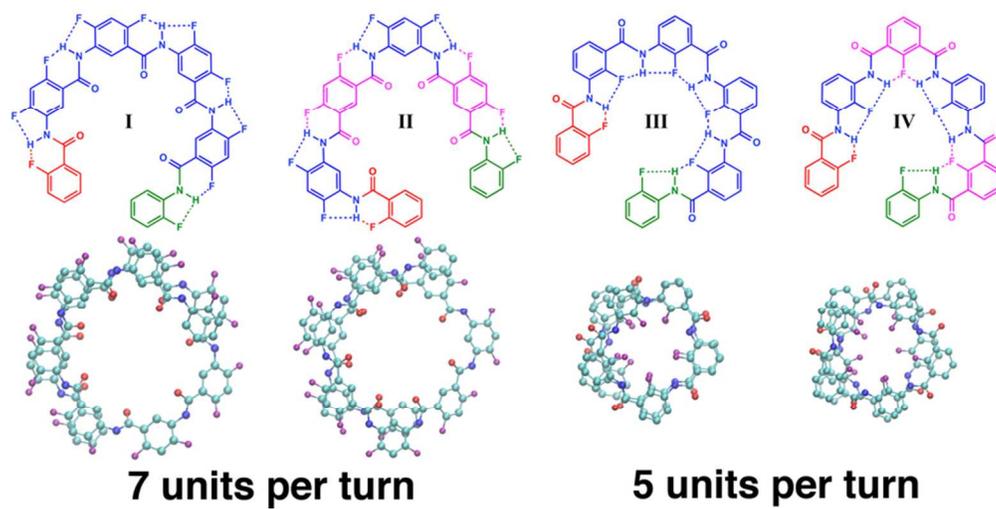
Snapshots from molecular dynamics simulations showcase how substituent positions and linkage types affect secondary structure properties of fluorobenzene based helical arylamides.
40x20mm (600 x 600 DPI)