

Folding of a Linear Array of α -Amino Acids within a Helical Aromatic Oligoamide Frame

Mayumi Kudo,^{†,‡,§} Victor Maurizot,^{‡,§} Brice Kauffmann,^{||,⊥,#} Aya Tanatani,^{*,†} and Ivan Huc^{*,‡,§}

[†]Department of Chemistry, Faculty of Science, Ochanomizu University, 2-1-1 Otsuka, Bunkyo-ku, Tokyo 112-8610, Japan [‡]Université de Bordeaux, CBMN (UMR 5248), Institut Européen de Chimie Biologie, 2 rue Escarpit, 33600 Pessac, France [§]CNRS, CBMN (UMR 5248), France

^{II}Université de Bordeaux, Institut Européen de Chimie Biologie (UMS 3033/US 001), 2 rue Escarpit, 33600 Pessac, France ^LCNRS, Institut Européen de Chimie Biologie (UMS 3033), 33600 Pessac, France

[#]INSERM, Institut Européen de Chimie Biologie (US 001), 33600 Pessac, France

(5) Supporting Information

ABSTRACT: Control of the spatial organization of proteinogenic side chains is critical for the development of protein mimics with selective recognition properties toward target protein surfaces. We present a novel methodology for producing a linear array of proteinogenic residues based on the incorporation of α -amino acids into sequences of rigid, helically folded oligoamides of 8-amino-2-quinolinecarboxylic acid (Q). When L-leucine (L) was alternated with dimer Q_2 , the resulting sequence adopted a right-handed helical conformation, as deduced in solution from the CD spectra of L- $(LQ_2)_n$ (n = 2, 4) and in the solid state from X-ray crystallographic analysis of (\pm) - $(LQ_2)_4$. Each LQ₂ segment spanned just one helix turn (pitch of 3.5 Å), and consequently, the four leucine side chains of $(LQ_2)_4$ formed a linear array. In solution, NMR analysis showed that both $L-(LQ_2)_2$ and $L-(LQ_2)_4$ exist as a mixture of two slowly equilibrating folded conformers, the proportion of which strongly varies with the solvent.

Protein tertiary structures and protein-protein recognition largely rest on intramolecular and intermolecular interactions between α -amino acid side chains projected at welldefined spatial positions from β -strands, turns, and helical subunits. Controlling the spatial organization of proteinogenic side chains has thus been a prime objective in peptide and protein mimics in order to fine-tune their ability to selfassemble or to interact with protein targets. For example, β and γ -peptides may adopt several types of helical conformations from which side chains are projected in a variety of arrays distinct from those of α -helical α -peptides.¹ The possibility of combining α -, β -, and γ -amino acids in the same helically folded sequence further increases the number of spatial arrays of side chains that can be created.² On the basis of these helix designs, artificial tertiary or quaternary structures³ and important steps toward protein surface recognition by non-natural sequences have been reported.⁴ Appending proteinogenic side chains to rigid scaffolds such as linear aryl oligomers (e.g., terphenyl groups and linear aryl oligoamides)⁵ or macrocycles⁶ in order to mimic protein epitopes has also been thoroughly investigated. In particular, strong emphasis has been given in recent years to linear arrays of side chains at defined intervals that can mimic the projection of *i*, *i* + 4, and *i* + 7 residues from one face of an α -helix.⁷ Here we present a novel approach for producing linear arrays of side chains via the incorporation of α -amino acids into rigid, helical aromatic oligoamide frames. The new folded motifs we have discovered possess facial polarity and provide a valuable alternative to currently known examples of peptide/protein mimics for protein surface recognition or as amphipathic structures to interact with bilayer membranes.

Aromatic oligoamide foldamers have emerged as an important class of folding oligomers characterized by high conformational stability and high predictability of their folded conformations.⁸ Oligoamides derived from 8-amino-2-quinolinecarboxylic acid (\mathbf{Q}) (Chart 1) fold into helices having 2.5

Chart 1. Structures of Q, LQ (4), and $(LQ_2)_n$ Oligomers 1–3



units per turn and a pitch of 3.5 Å, as demonstrated both in the solid state and in solution.⁹ These helices are extremely stable in essentially any solvent.^{10,9} For example, they show no denaturation at 120 °C in dimethyl sulfoxide. Their robustness allows them to accommodate a number of more flexible

 Received:
 May 15, 2013

 Published:
 June 13, 2013

Journal of the American Chemical Society

aliphatic units that a priori are not prone to folding, to which they dictate folding behavior compatible with the aromatic oligoamide Q_n helix.¹¹ This contrasts with other examples of hybrid aromatic–aliphatic sequences, which often display novel unexpected folding behaviors.¹²

 α -Amino acids have no apparent feature that would make them compatible with helically folded Q_n oligoamides. Nevertheless, we were interested in exploring the folding behavior of combinations thereof that would allow us to exploit the robustness of Q_{μ} conformations and the wide range of α amino acid residues readily available from commercial sources. Given that a Q_2 dimer spans 0.8 helix turn, we speculated that an additional α -amino acid such as leucine (L) would allow an $\boldsymbol{L}\boldsymbol{Q}_2$ segment 13 to span just about one turn and that repeating this segment in the same sequence would potentially bring the leucine residues into close proximity. In the following, we report the synthesis and structural characterization of $(LQ_2)_n$ oligomers and the discovery that they adopt helically folded conformations from which the leucine residues are projected in a linear array from one face of the helix. The aromatic amide helix provides a rigid frame into which a certain number of α amino acids may be inserted and from which linear arrays of side chains may be projected for molecular recognition purposes or to build amphipathic structures.

The synthesis of $(LQ_2)_n$ oligomers¹³ is presented in detail in the Supporting Information (SI). It makes use of optimized monomer and oligomer synthetic procedures.¹⁴ In short, the main-chain terminal aliphatic amine was protected by a Boc group. The main-chain terminal acid was protected as a 2-(trimethylsilyl)ethyl (TMSE) ester, which could be removed in the presence of fluoride and thus avoid basic conditions that might racemize leucine α -carbons. The direct coupling of Boc-L-leucine to the free amine of a quinoline dimer to give $LQ_2(1)$ (Chart 1) failed in our hands. Even with acid fluoride activation of leucine, it required long heating times in the presence of N,N-diisopropylethylamine that eventually led to racemization of the leucine α -carbon (see the SI). On the other hand, the acid fluoride of Boc-L-leucine could be coupled to an 8aminoquinoline monomer to give LQ(4) in 93% yield without any racemization (Chart 1). Cleavage of the ester of 4, activation of the resulting acid as an acid chloride under neutral conditions, and coupling to another 8-aminoquinoline gave 1. Iterative deprotections of the N- and C-termini, activation of the terminal acid with HBTU, and coupling to the terminal aliphatic amine then allowed $(LQ_2)_2$ (2) and $(LQ_2)_4$ (3) to be prepared in a convergent fashion. For the purpose of racemic crystallographic investigations (see below), the synthesis was carried out twice, in the L series and in the D series.

Preliminary data concerning the conformational behaviors of these new aliphatic—aromatic hybrids in solution were collected using circular dichroism (CD) spectroscopy (Figure 1). Similar behaviors were observed in CHCl₃, acetonitrile, and acetone solutions. While L-4 showed no CD signal in the 200–400 nm range, L-1 showed distinct negative bands at 252.9 nm ($\Delta \varepsilon =$ -48.3 cm² mmol⁻¹) and 370.0 nm ($\Delta \varepsilon =$ -19.2 cm² mmol⁻¹) in CHCl₃. According to previous assignments, the negative band at 370 nm is suggestive of a left-handed helix twist.^{15a} L-2 and L-3 featured intense CD bands characteristic of a folded structure, but their signs were reversed from that of L-1, indicating the opposite handedness (right-handed). The CD band intensities per LQ₂ repeat unit were similar for L-2 and L-3, suggesting that the extent of folding was similar for these two species.





Figure 1. CD spectra of 1-4 in CHCl₃, CH₃CN, and acetone at 25 °C. The horizontal scales have been set to regions above the absorption ranges of the solvents.

Attempts to grow crystals of 1–3 as single L-enantiomers that would be suitable for X-ray crystallographic analysis all failed. Our experience with helical aromatic oligoamides is that racemic or quasi-racemic crystals grow much more readily,¹⁵ as is the case for other peptides and small proteins.¹⁶ This proved to be valid for the $(LQ_2)_n$ oligomers as well, and the structures of *rac*-1 and *rac*-3 in the solid state could both be solved in the centrosymmetric space group $P\overline{1}$. The structure of 1 (Figure 2a) revealed that the L enantiomer adopts a left-handed helical conformation, consistent with the sign of the CD band observed in solution (Figure 1). The helix is stabilized by a hydrogen bond ($d_{N-O} = 2.9$ Å) between the N-terminal NH proton and the C-terminal carbonyl oxygen atom, which are located almost exactly above each other when viewed along the



Figure 2. Views of (a-c) the crystal structures of (a) the left-handed helix of L-1 and (b, c) the right-handed helix of L-3 and (d, e) a molecular model of an alternative structure of 3. Leucine N atoms and side chains are shown in blue and gold, respectively. Some quinoline or carbamate carbonyl O atoms are shown in red. Some key hydrogen bonds are shown as dashed green lines. Included solvent molecules, isobutoxy chains, TMSE groups, *tert*-butyl groups, and H atoms have been omitted for clarity.

helix axis, confirming the initial assumption that the LQ₂ motif would span one helix turn. This hydrogen bond sets the Lleucine with $\phi = -117^{\circ}$ and $\psi = -5^{\circ}$, resulting in an almost perfect anti conformation of the C^{α}H and NH bonds (179^{\circ}). The N-terminal carbamate CONH plane is parallel to the helix axis.

In contrast, L-3 adopts a right-handed helical conformation, also in agreement with the sign of its CD band. The leucine amide NHs (but not the carbamate NH) point toward the helix axis and form hydrogen bonds to the endocyclic N atoms of the neighboring i - 1 quinoline units. Each of the four LQ₂ segments accounts for about one helix turn (pitch of 3.5 Å), resulting in a linear array of the leucine side chains (see the top view in Figure 2b), which protrude in an almost perpendicular fashion from the helix axis. The $C^{\alpha}-C^{\alpha}$ distances between consecutive leucines are 4.8, 4.7, and 4.2 Å starting from the Nterminus; these are smaller than the distances between the i, i + 4, and i + 7 residues of α -helices (6.3 Å on average). Starting from the N terminus, the L-leucines in L-3 have ϕ values of -133, -143, -140, and -155° and ψ values of -5, 26, 40, and 35°. Such high negative ϕ values are common in peptides, as they can be found both in β -sheets and α -helices; the small positive ψ values fall in an allowed (though not favored) area of the Ramachandran plot.

NMR studies were carried out to investigate the conformations in solution (Figure 3; also see the SI). In



Figure 3. Parts of the 300 MHz ¹H NMR spectra of **1** and the 400 MHz ¹H NMR spectra of **2** and **3** showing aromatic amide resonances (10–12 ppm), some aliphatic amide and aromatic proton resonances (8–9 pmm), and TMS resonances (–0.5 to 0 ppm): (a) **1** at 25 °C in CDCl₃; (b–d) **2** at 45, 25, and 5 °C in CDCl₃; (e) **2** at 5 °C in CDCl₃; (f) **3** at 25 °C in CDCl₃. The stars indicate two leucine NH protons having very different ${}^{3}J_{NH-C^{\circ}H}$ coupling constants.

going from 1 to 3, the ¹H NMR spectra showed upfield shifts of most of the signals, including those of the terminal groups, indicating increasing and additive ring-current effects associated with $\pi-\pi$ stacking. For example, the signals of the terminal TMS groups in 1, 2, and 3 appeared at 0.04, -0.16, and -0.30 ppm, respectively, in CDCl₃ at 25 °C. Such a chain-length dependence is typical of folding phenomena.¹⁷ While the spectrum of 1 in CDCl₃ was sharp at 25 °C, the spectra of 2 and 3 showed some broad signals that sharpened upon heating to 45 °C. Cooling caused the signals of 2 and 3 first to broaden further and then to split into two sets of sharp signals near 5 °C. This behavior indicates the coexistence of two well-defined

species that equilibrate slowly on the NMR time scale. A possible aggregated state was ruled out, as the proportions of these two species did not depend upon concentration. However, the proportions did show a strong solvent dependence. In the case of **2**, proportions of 53:47, 88:12, and 85:15 were measured at 5 °C in CDCl₃, CD₃CN, and acetone- d_8 , respectively. Spectra in solvent mixtures confirmed that the major species was the same in all three cases. This large variation of the proportions contrasts with the CD bands, which had the same intensity in all three solvents, indicating that the two species observed by NMR could not be assigned to two diastereomeric *P* and *M* helices and that the induction of handedness was probably quantitative in the case of **2** and **3**.¹⁵

The above data point to the existence of two well-defined nonaggregated states having the same handedness. Direct full NMR assignments and structure elucidation were hampered by broadness and the fact that exchange phenomena may be difficult to distinguish from nuclear Overhauser effect correlations. The sharp spectrum of 2 at 5 $^\circ C$ allowed us to record correlation spectroscopy (COSY) spectra, from which the NHs of the central leucine unit were assigned to two signals at 8.98 and 8.69 ppm for the major and minor species, respectively. The coupling constants were measured to be ${}^{3}J_{\text{NH}-\text{C}^{\alpha}\text{H}}$ = 6.9 and 10.5 Hz for the signals at 8.98 and 8.69 ppm, respectively. These values indicate quite different conformations of the leucine unit. The large coupling constant above 10 Hz most likely matches the conformation observed in the crystal structure of 3. Indeed, applying the Karplus equation to the ${}^{3}J_{\rm NH-C^{a}H}$ couplings indicated ϕ values close to -120° and a $C^{\alpha}H-NH$ dihedral angle close to 180° (see the SI). In contrast, the coupling constant at 6.9 Hz may correspond to various ϕ values near -160, -80, or +60°.

The examination of molecular models led to the proposal of an alternative structure of 3 different from its structure in the solid state (Figure 2d,e). This structure was inspired by the structure of 1 in the solid state (Figure 2a), as it also involves intramolecular hydrogen bonds between the leucine NH protons of unit *i* and the amide carbonyl group of quinoline unit i + 2 instead of the quinoline endocyclic nitrogen atom of unit i - 1. When this pattern is repeated along the four leucines of 3, the corresponding amide functions tilt to planes parallel to the helix axis and form a linear array of hydrogen bonds reminiscent of α -helices. This is accommodated without a change in the helix pitch (3.5 Å) by a slight increase in the helix diameter and a tilt of the array of leucines with respect to the helix axis. This conformation is very stable in molecular dynamics simulations (see the SI), making it a plausible hypothetical candidate for one of the species observed in solution. Indeed, in this conformation, the leucine C^{α} -H and N–H are eclipsed, leading to ϕ values of +60°, consistent with the NMR observations mentioned above.

In summary, sequences combining the α -amino acid leucine (L) and 8-amino-2-quinolinecarboxylic acid (Q) units in the particular arrangement (LQQ)_n give rise to folded conformations in which linear arrays of proteinogenic side chains are produced. The next step consists in bringing such sequences into water, where aromatic oligoamide helical folding is dramatically enhanced^{10b,18} and where stable patterns might be expected to form. Research in this direction is in progress and will be reported in due course.

Journal of the American Chemical Society

ASSOCIATED CONTENT

S Supporting Information

Experimental procedures, full characterization of new compounds, crystallographic data (CIF), and detailed NMR and CD investigations. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

tanatani.aya@ocha.ac.jp; i.huc@iecb.u-bordeaux.fr

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the International Training Program of JSPS (predoctoral fellowship to M.K.). We thank Dr. Shigeru Ito for assistance with interpretation of NMR spectra.

REFERENCES

(1) (a) Appella, D. H.; Christianson, L. A.; Karle, I. L.; Powell, D. R.; Gellman, S. H. *J. Am. Chem. Soc.* **1996**, *118*, 13071. (b) Seebach, D.; Overhand, M.; Kühnle, F. N. M.; Martinoni, B.; Oberer, L.; Hommel, U.; Widmer, H. *Helv. Chim. Acta* **1996**, *79*, 913. (c) Guo, L.; Chi, Y.; Almeida, A. M.; Guzei, I. A.; Parker, B. K.; Gellman, S. H. J. Am. Chem. Soc. **2009**, *131*, 16018.

(2) (a) Vasudev, P. G.; Chatterjee, S.; Shamala, N.; Balaram, P. Chem. Rev. 2011, 111, 657. (b) Hayen, A.; Schmitt, M. A.; Ngassa, F. N.; Thomasson, K. A.; Gellman, S. H. Angew. Chem., Int. Ed. 2004, 43, 505. (c) Sharma, G. V.; Reddy, K. R.; Krishna, P. R.; Sankar, A. R.; Narsimulu, K.; Kumar, S. K.; Jayaprakash, P.; Jagannadh, B.; Kunwar, A. C. J. Am. Chem. Soc. 2003, 125, 13670. (d) Mándity, I. M.; Wéber, E.; Martinek, T. A.; Olajos, G.; Tóth, G. K.; Vass, E.; Fülöp, F. Angew. Chem., Int. Ed. 2009, 48, 2171. (e) Hetényi, A.; Tóth, G. K.; Somlai, C.; Vass, E.; Martinek, T. A.; Fülöp, F. Chem.—Eur. J. 2009, 15, 10736. (f) Claudon, P.; Violette, A.; Lamour, K.; Decossas, M.; Fournel, S.; Heurtault, B.; Godet, J.; Mély, Y.; Jamart-Grégoire, B.; Averlant-Petit, M.-C.; Briand, J.-P.; Duportail, G.; Monteil, H.; Guichard, G. Angew. Chem., Int. Ed. 2010, 49, 333. (g) Guo, L.; Almeida, A. M.; Zhang, W.; Reidenbach, A. G.; Choi, S. H.; Guzei, I. A.; Gellman, S. H. J. Am. Chem. Soc. 2010, 132, 7868.

(3) (a) Daniels, D. S.; Petersson, E. J.; Qiu, J. X.; Schepartz, A. J. Am. Chem. Soc. 2007, 129, 1532. (b) Horne, W. S.; Price, J. L.; Keck, J. L.; Gellman, S. H. J. Am. Chem. Soc. 2007, 129, 4178.

(4) (a) Michel, J.; Harker, E. A.; Tirado-Rives, J.; Jorgensen, W. L.; Schepartz, A. J. Am. Chem. Soc. 2009, 131, 6356. (b) Sadowsky, J. D.; Schmitt, M. A.; Lee, H.-S.; Umezawa, N.; Wang, S.; Tomita, Y.; Gellman, S. H. J. Am. Chem. Soc. 2005, 127, 11966. (c) Boersma, M. D.; Haase, H. S.; Peterson-Kaufman, K. J.; Lee, E. F.; Clarke, O. B.; Colman, P. M.; Smith, B. J.; Horne, W. S.; Fairlie, W. D.; Gellman, S. H. J. Am. Chem. Soc. 2012, 134, 315. (d) Robinson, J. A. ChemBioChem 2009, 10, 971. (e) Peczuh, M. W.; Hamilton, A. D. Chem. Rev. 2000, 100, 2479. (f) Kritzer, J. A.; Lear, J. A.; Hodsdon, M. E.; Schepartz, A. J. Am. Chem. Soc. 2004, 126, 9468.

(5) (a) Kutzki, O.; Park, H. S.; Ernst, J. T.; Orner, B. P.; Yin, H.; Hamilton, A. D. J. Am. Chem. Soc. 2002, 124, 11838. (b) Cummings, C. G.; Hamilton, A. D. Curr. Opin. Chem. Biol. 2010, 14, 341.
(c) Plante, J. P.; Burnley, T.; Malkova, B.; Webb, M. E.; Warriner, S. L.; Edwards, T. A.; Wilson, A. J. Chem. Commun. 2009, 5091. (d) Yin, H.; Lee, G.-i.; Sedey, K. A.; Rodriguez, J. M.; Wang, H.-G.; Sebti, S. M.; Hamilton, A. D. J. Am. Chem. Soc. 2005, 127, 5463.

(6) (a) Park, H. S.; Lin, Q.; Hamilton, A. D. *J. Am. Chem. Soc.* **1999**, *121*, 8. (b) Fasan, R.; Dias, R. L. A.; Moehle, K.; Zerbe, O.; Obrecht, D.; Mittl, P. R. E.; Grütter, M. G.; Robinson, J. A. *ChemBioChem* **2006**, *7*, 515.

(7) Wilson, A. J. Chem. Soc. Rev. 2009, 38, 3289.

(8) (a) Huc, I. Eur. J. Org. Chem. 2004, 17. (b) Zhang, D.-W.; Zhao, X.; Hou, J.-L.; Li, Z.-T. Chem. Rev. 2012, 112, 5271.

(9) (a) Jiang, H.; Léger, J.-M.; Huc, I. J. Am. Chem. Soc. 2003, 125, 3448. (b) Jiang, H.; Léger, J.-M.; Dolain, C.; Guionneau, P.; Huc, I. Tetrahedron 2003, 59, 8365. (c) Dolain, C.; Grélard, A.; Laguerre, M.; Jiang, H.; Maurizot, V.; Huc, I. Chem.—Eur. J. 2005, 11, 6135.

(10) (a) Delsuc, N.; Kawanami, T.; Lefeuvre, J.; Shundo, A.; Ihara, H.; Takafuji, M.; Huc, I. *ChemPhysChem* **2008**, *9*, 1882. (b) Qi, T.; Maurizot, V.; Noguchi, H.; Charoenraks, T.; Kauffmann, B.; Takafuji, M.; Ihara, H.; Huc, I. *Chem. Commun.* **2012**, *48*, 6337.

(11) (a) Sánchez-García, D.; Kauffmann, B.; Kawanami, T.; Ihara, H.; Takafuji, M.; Delville, M.-H.; Huc, I. *J. Am. Chem. Soc.* 2009, 131, 8642. (b) Baptiste, B.; Douat-Casassus, C.; Laxmi-Reddy, K.; Godde, F.; Huc, I. *J. Org. Chem.* 2010, 75, 7175. (c) Delsuc, N.; Poniman, L.; Léger, J.-M.; Huc, I. *Tetrahedron* 2012, 68, 4464.

(12) (a) Prabhakaran, P.; Kale, S. S.; Puranik, V. G.; Rajamohanan, P. R.; Chetina, O.; Howard, J. A.; Hofmann, H. J.; Sanjayan, G. J. J. Am. Chem. Soc. 2008, 130, 17743. (b) Srinivas, D.; Gonnade, R.; Ravindranathan, S.; Sanjayan, G. J. J. Org. Chem. 2007, 72, 7022.
(c) Roy, A.; Prabhakaran, P.; Baruah, P. K.; Sanjayan, G. J. Chem. Commun. 2011, 47, 11593.

(13) Throughout the manuscript, sequences abbreviated $(LQ_2)_n$ possess a Boc group at the N-terminus and a TMSE ester at the C-terminus.

(14) Qi, T.; Deschrijver, T.; Huc, I. Nat. Protoc. 2013, 8, 693.

(15) (a) Dolain, C.; Jiang, H.; Léger, J.-M.; Guionneau, P.; Huc, I. J. Am. Chem. Soc. **2005**, 127, 12943. (b) Maurizot, V.; Dolain, C.; Leydet, Y.; Léger, J.-M.; Guionneau, P.; Huc, I. J. Am. Chem. Soc. **2004**, 126, 10049. (c) Delsuc, N.; Massip, S.; Léger, J.-M.; Kauffmann, B.; Huc, I. J. Am. Chem. Soc. **2011**, 133, 3165. (d) Ferrand, Y.; Kendhale, A. M.; Kauffmann, B.; Grélard, A.; Marie, C.; Blot, V.; Pipelier, M.; Dubreuil, D.; Huc, I. J. Am. Chem. Soc. **2010**, 132, 7858.

(16) (a) Toniolo, C.; Peggion, C.; Crisma, M.; Formaggio, F.; Shui, X.; Eggleston, D. S. *Nat. Struct. Biol.* **1994**, *1*, 908. (b) Pentelute, B. L.; Gates, Z. P.; Tereshko, V.; Dashnau, J. L.; Vanderkooi, J. M.; Kossiakoff, A.; Kent, S. B. H. *J. Am. Chem. Soc.* **2008**, *130*, 9695.

(17) Stone, M. T.; Heemstra, J. M.; Moore, J. S. Acc. Chem. Res. 2006, 39, 11.

(18) Gillies, E. R.; Deiss, F.; Staedel, C.; Schmitter, J.-M.; Huc, I. Angew. Chem., Int. Ed. 2007, 46, 4081.