A Peptoid with Extended Shape in Water

Jumpei Morimoto[1,*], Yasuhiro Fukuda[1,*], Takumu Watanabe[1], Daisuke Kuroda[1,2], Kouhei Tsumoto[1,2,3], and Shinsuke Sando[1,2]

[1]Department of Chemistry and Biotechnology, Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan. [2]Department of Bioengineering, Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan. [3]Medical Proteomics Laboratory, The Institute of Medical Science, The University of Tokyo, Shirokanedai, Minato-ku, Tokyo, 108–8639, Japan
*These authors contributed equally to this work.

Abstract

"Peptoids" was proposed, over decades ago, as a term describing analogs of peptides that exhibit better physicochemical and pharmacokinetic properties than peptides. Oligo-(*N*-substituted glycines) (oligo-NSG) was previously proposed as a peptoid due to its high proteolytic resistance and membrane permeability. However, oligo-NSG is conformationally flexible and is difficult to achieve a defined shape in water. This conformational flexibility is severely limiting biological application of oligo-NSG. Here, we propose oligo-(*N*-substituted alanines) (oligo-NSA) as a new peptoid that forms a defined shape in water. A synthetic method established in this study enabled the first isolation and conformational study of optically pure oligo-NSA. Computational simulations, crystallographic studies and spectroscopic analysis demonstrated the well-defined extended shape of oligo-NSA realized by backbone steric effects. The new class of peptoid achieves the constrained conformation without any assistance of *N*-substituents and serves as an ideal scaffold for displaying functional groups in well-defined three-dimensional space, which leads to effective biomolecular recognition.

Introduction

"Peptoids"[1,2] was coined for analogs of peptides that exhibit better physicochemical and pharmacokinetic properties than peptide (**Figure 1a**). Oligo(*N*-substituted glycines) (oligo-NSG) was proposed as an example of peptoid in 1992[3] because the oligomer is highly resistant to proteolytic degradation[4] and exhibits high membrane permeability[5] due to its *N*-substituted amide backbone. Besides, submonomer synthetic methods of oligo-NSG allow introduction of a diverse set of proteinogenic and nonproteinogenic functional groups as *N*-substituents.[6,7] However, backbone of the NSG-type peptoid is intrinsically flexible, therefore, peptoid that exhibits high affinity to biomolecules has not been reported except for rare exceptions.[8] To overcome the limitation, extensive researches have been devoted to constrain the conformation of peptoids by using elaborate *N*-substituents but such conformational regulation generally requires introduction of multiple bulky hydrophobic *N*-substituents that severely compromise water solubility[9–14]. Introduction of charged groups is effective to compensate the water solubility of these peptoids[15–18] but these restrictions reduce the design flexibility of oligo-NSGs.

Introduction of methyl substituents on backbone α-carbon of oligo-NSG to make the molecule as oligo-(*N*-substituted alanines) (oligo-NSA) was previously proposed to constrain the conformation of peptoid via the local steric effects on backbone (**Figure 1a**).[19,20] The constrained conformation was reasoned from an analogy of *N*-methylated peptides. More specifically, the introduced methyl substituent would produce steric repulsion, known as pseudo-1,3-allylic strain[21,22], with the carbonyl oxygen of the preceding residue and the *N*-substituent of the following residue and restrict rotation about φ and ψ angles (**Figure 1b**). Moreover, steric repulsion between two methyl substituents on neighboring residues would destabilize cis conformation of the intervening amide bond and bias the amide bond to trans; ω angles are restricted to the value around 180° (**Figure 1c**). Because it does not require introduction of bulky *N*-substituents but require introduction of only a small methyl substituent, it is an ideal strategy to realize "peptoid" that forms a defined shape in biological environment and displays various functional groups as *N*-substituents into well-defined three-dimensional space. Despite the potential utility of oligo-NSA, there have been no reports of optically pure oligo-NSAs possessing *N*-substituents other than methyl due to the synthetic difficulty of such oligomers. Thus, the validity of the constrained conformation of oligo-NSA remains to be examined and establishment of synthetic scheme for such a useful oligomer has been awaited.

We here established a synthetic method of oligo-NSA and isolated optically pure oligo-NSAs for the first time. With the synthesized olio-NSAs, conformational studies of oligo-NSAs were conducted and the studies revealed that oligo-NSA achieves a defined shape in water and serves as a scaffold of an extended shape displaying functional groups in well-defined three-dimensional space in water.

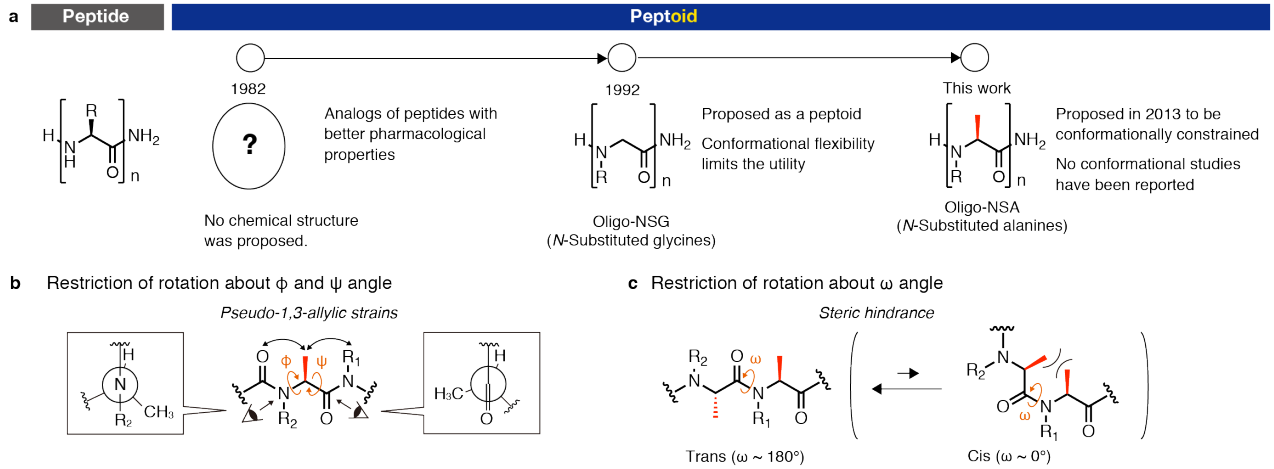


Figure 1 | Chemical structure and proposed conformational regulation by steric effects of oligo-NSA. (**a**) A brief overview of the history of peptoid. (**b**) Bond rotations about φ and ψ angles of oligo-NSA are restricted due to the pseudo-1,3-allylic strains. (**c**) ω angle of oligo-NSA is biased to ~180° due to the steric repulsion between backbone structures.

Quantum mechanical studies

First, we performed computational simulations to predict how the pseudo-1,3-allylic strain would constrain the conformation of peptoid and how the conformation of oligo-NSA would look like. A minimal component of oligo-NSG, i.e. acetyl-*N*-methylglycine dimethylamide, and oligo-NSA, i.e. acetyl-*N*-methylalanine dimethylamide, were subjected to quantum mechanical (QM) calculations to generate Ramachandran-type energy diagrams over φ and ψ angles. The diagram of oligo-NSG had a relatively large region within 10 kcal/mol from the lowest energy point (**Figure 2a**), which is consistent with the previous reports of similar calculations.[23,24] In contrast, the diagram of oligo-NSA had a narrow region with the lowest energy at around (φ, ψ) = (–120°, 90°) and the surrounding region was predominantly favored over other regions (**Figure 2b**). To understand how the *N*-substituents will be oriented on oligo-NSA, we have also examined preferred χ angles of acetyl-*N*-ethylalanine dimethylamide with fixed φ and ψ angles of the lowest energy point; (φ, ψ) = (–120°, 90°). As a result, 100° and –100° were determined to be the χ angles of the lowest energy (**Figure S1**). Based on this χ scan, energy diagrams of NSA over φ and ψ angles were re-examined using the χ angle of 100° and –100° for acetyl-*N*-ethylalanine dimethylamide as a minimal model of oligo-NSA with *N*-substituents larger than methyl. For χ angle of 100°, the lowest energy angles were slightly shifted and became (φ, ψ) = (–105°, 105°) but, overall for both the χ angles, the allowed region was similar to the above-mentioned landscape of acetyl-*N*-methylalanine dimethylamide (**Figure 2c and d**), suggesting that the structure of *N*-substituent does not largely affect to the backbone conformation. These results support the hypothesis that introduction of a methyl substituent on α-carbon of NSG would locally restrict the bond rotations about all the backbone dihedral angles.

Based on the QM calculation of the monomeric NSA structure, we generated model structures of oligo-NSA. The models were built with the dihedral angles (χ, φ, ψ, ω) of either (100°, –105°, 105°, 180°) or (–100°, –120°, 90°, 180°) and optimized with QM calculations with an implicit water model (**Figure 2e and f**). With either set of dihedral angles, the optimized structure of the oligo-NSA resulted in an extended shape. The conformation with χ angle of –100° was 1.4 kcal/mol more stable than the conformation with χ angle of 100°, indicating that χ angle form –100° a little more preferably in an oligomer. The extended shape of oligo-NSA is realized by the iterative regulation of backbone dihedral angles driven by the steric effects. This backbone-dictated shape would be an ideal scaffold for biomolecular recognition because the backbone conformation does not require elaborate *N*-substituents and functional groups introduced as *N*-substituents would be displayed in well-distributed space without the oligomer collapsing into a more compact shape.

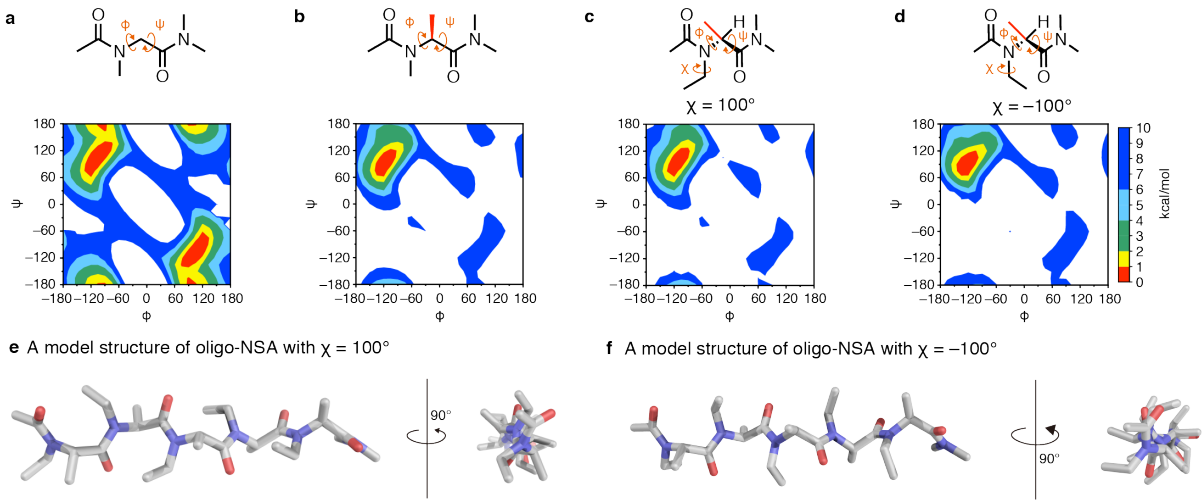
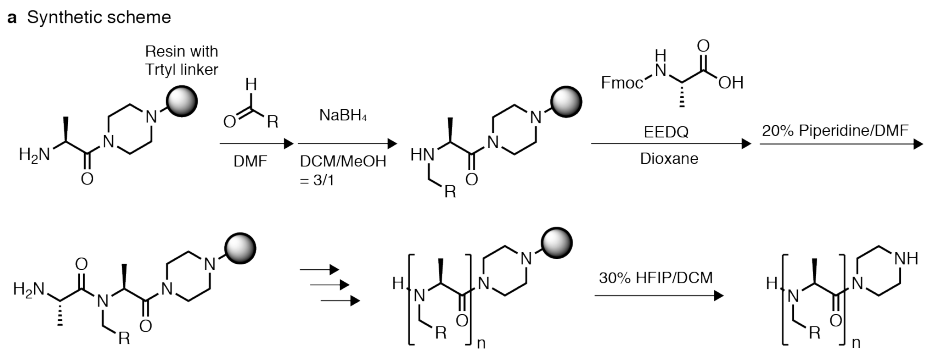


Figure 2 | Ramachandran-type energy diagrams of NSG and NSA. (**a**) An energy diagram of acetyl-*N*-methylglycine dimethylamide as a minimal model of oligo-NSG. (**b**) An energy diagram of acetyl-*N*-methylalanine dimethylamide as a minimal model of oligo-NSA. (**c, d**) An energy diagram of acetyl-*N*-ethylalanine dimethylamide as a minimal model of oligo-NSA with *N*-substituents larger than methyl. χ angle was fixed to 100° for **c** and –100° for **d** based on the result of χ scan (**Figure S1**). For these diagrams, regions that have energies above 10 kcal/mol from the lowest energy point were colored in white. (**e**) A model structure of acetyl-*N*-ethylalanine pentamer that was optimized from an initial conformation of (χ, φ, ψ, ω) = (100°, –105°, 105°, 180°) with a QM calculation. (**f**) A model structure of acetyl-*N*-ethylalanine pentamer that was optimized from an initial conformation of (χ, φ, ψ, ω) = (–100°, –120°, 90°, 180°) with a QM calculation.

Synthesis

To experimentally validate the extended conformation of oligo-NSA, optically pure oligomers need to be isolated in multi-milligram quantity. Although there are reports about peptides or oligo-NSGs with non-consecutive NSA residues, there have been no reports for oligomers with consecutive NSA residues with *N*-substituents other than methyl. The extended shape of oligo-NSA would be realized only realized by such oligomers with consecutive NSA residues, therefore, we first established a synthetic method for optically pure oligo-NSAs. The most promising synthetic method was the submonomer synthetic method using chiral 2-bromopropionic acid and primary amine (**Figure S2a**) that was proposed by Zuckerman and Kodadek[19,25] because this method does not require prior preparation of individual *N*-substituted alanines but only requires commercially available simple submonomers. However, due to the racemization prone nature of 2-bromopropionic acid[26], efficient synthesis of optically pure oligo-NSAs using the method was unsuccessful (**Figure S2**). Therefore, we turned to another synthetic method of oligo-NSAs reported by Pels and Kodadek that utilizes Fmoc-Ala-OH and aldehydes as submonomers[7] (**Figure S3a**). This method also only requires commercially available submonomers and equally attractive to the abovementioned submonomer synthetic method. However, in their report, consecutive introduction of NSA residues was avoided due to the inefficient coupling reaction of Fmoc-alanine on *N*-substituted alanine terminus. Therefore, we investigated various coupling conditions for Fmoc-alanine on *N*-substituted alanine terminus to achieve introduction of consecutive NSA residues (**Figure S3b**). Among the tested conditions, coupling reaction using 1-ethoxycarbonyl-2-ethoxy-1,2-dihydroquinoline (EEDQ) in dioxane gave the highest coupling yield (**Figure S3c** and **Figure S4**). Double coupling at 60 °C for 3 h each or at 100 °C by microwave heating for 1 h each using EEDQ gave a quantitative conversion with no detectable amount of racemization (**Figure S3c and d**). Encouraged by the quantitative conversion for the challenging acylation reaction, 1–5 mer of NSAs were synthesized on resin using the optimized conditions (**Figure 3a**). Note that trityl linker was recruited to allow cleavage of the synthesized oligomers from resin with a mild acidic condition (30% hexafluoroisopropanol in dichloromethane) due to the highly acid labile nature of NSAs[27,28] and piperazine spacer was inserted between trityl linker and NSA oligomers to prevent diketopiperazine formation during synthesis. All the 1–5 mer of NSAs were obtained with decent yields (**Figure 3b, S5–8**). There have been reports about oligo-NSGs containing skipped NSA residues[7,19,29–33] but this is the first report of isolation of optically pure oligo-NSAs containing consecutive NSA residues with *N*-substituents other than methyl.



Compound	n	Yield (%) HPLC	Yield (%) Isolated
1	1	82	80
2	2	85	71
3	3	72	54
4	4	83	55
5	5	58	47

Figure 3 | Submonomeric synthesis of oligo-NSA using Fmoc-Ala-OH and aldehyde. (a) Synthetic scheme of oligo-NSA. Resin with trityl linker was utilized to allow cleavage of the synthesized compounds from resin using mild acidic conditions. The piperazine spacer was introduced between the trityl linker and oligo-NSAs to prevent diketopiperazine formation during synthesis. **(b)** Yields of *N*-isobutylalanine 1–5 mer (**1**–**5**). HPLC chromatograms of crude and purified products are listed on **Figure S5** and **S6**. For monomer (**1**), 1H and 13C NMR spectra are shown on **Figure S7** and **S8**.

X-ray crystallographic study and molecular dynamics simulations

With the optimized synthetic method, several oligo-NSAs were synthesized and X-ray crystallographic analysis was conducted to experimentally reveal a preferred conformation of oligo-NSA. Several oligo-NSAs were synthesized and applied to crystallization. Among various conditions tested for the synthesized oligomers, crystals of *N*-benzylalanine pentamer (**Figure 4a**, compound **6**) were grown via slow evaporation from a co-solvent of hexane and dichloromethane. The structure was solved by X-ray crystallographic analysis (**Figure 4b** and **c**, and **Figure S9**). In the crystal, the oligomer exhibited an extended shape that is consistent with the predicted steric effects and QM calculations (**Figure 4c**). First, all the amide bonds were trans as expected from the steric repulsion of backbone structures. Second, pseudo-1,3-allylic strains were seen at all the φ and ψ angles and, as a consequence, dihedral angles of all the residues were restricted to the region within 2 kcal/mol from the lowest energy point on the energy diagram generated by the QM calculations (**Figure 4d**). To assess how the crystal structure relates to the above-described model structure, the crystal structure was overlayed with one of the model structure shown in **Figure 2f** that gave a lower energy than the other model structure shown in **Figure 2e**. Overall, the backbone conformation of the crystal structure and the model structure almost completely overlapped with each other except for the *N*-terminal residue that is inevitably different because the *N*-terminal amine of the crystalized oligomer is not acylated and do not experience pseudo-1,3-allylic strain (**Figure 4e**). As seen in the model structure, also in the crystal, all the *N*-substituents were well separated on the oligomer and no obvious contacts with backbone or other *N*-substituents were observed, suggesting that the extended shape of oligo-NSA is dictated solely by backbone steric effects and independently from *N*-substituents. This makes the oligo-NSA as an ideal scaffold for displaying functional groups into well-dispersed three-dimensional space.

To obtain insights on how much the conformation of oligo-NSA observed in the crystal is stable in aqueous solution, molecular dynamics (MD) calculations of the *N*-benzylalanine pentamer were conducted using the crystal structure as the initial conformation. The calculation was performed for 500 ns at 298 K with five trials using the CHARMM36m force field and the TIP3P water model. As a result, φ and ψ angles were in the region within 4 kcal/mol from the lowest energy point on the energy diagram generated from the QM calculations for 95% of the simulated time (**Figure 4f**). ω angles were also kept at around 180° for all the simulated time (**Figure S10**) although this does not necessarily support the stability of trans amide bond because the kinetics of trans/cis isomerization *N*-substituted amide is much slower than the simulated time.[34] Therefore, the ω angles need to be also evaluated spectroscopically. (See the NMR analysis below.) As a consequence of the restricted rotations about these backbone dihedral angles, the oligomer well sustained the extended shape with a root mean squared deviation (RMSD) value of the α-carbons of the oligomer of only 0.5–0.9 Å on average, when compared to the initial crystal structure, for all the five runs (**Figure 4g**). As an exception, the RMSD value of the third run increased to ~2 Å at around 400 ns. The conformations

observed during the last 100 ns of the third run are little bent from the initial extended conformation due to transitions of φ angles from around at –120° to around at 70°. However, the dihedral angles seen in the bent conformations correspond to a high-energy region on the Ramachandran-type plot (the upper right blue region of **Figure 2b**), therefore the conformations are presumably a minor population. In addition, χ angles of the oligomers were also restricted to around –100° and 100° that were calculated to be low in energy in the QM calculations (**Figure S11**). In contrast, when the same MD simulations were run for the oligomer of NSG backbone using the same backbone conformation as the initial structure of oligo-NSA, φ and ψ angles of all the residues were distributed in much larger ranges (**Figure S12**) and the RMSD value was also much larger; 1.7 Å on average (**Figure 4h**). Interestingly, when an oligomer with alternate NSA and NSG residues were subjected to the same calculation, dihedral angles of NSA residues were constrained but those of NSG residues were not constrained (**Figure S13a**), which led to distortion of the initial extended shape and a large conformational fluctuation (**Figure S13b**). This result agrees with the assumption that the extended shape is only realized by oligomers with consecutive NSA residues with no intervening NSG residues. This is consistent with the previous report that oligo-NSGs with alternate NSA residues are intrinsically disordered.[30]

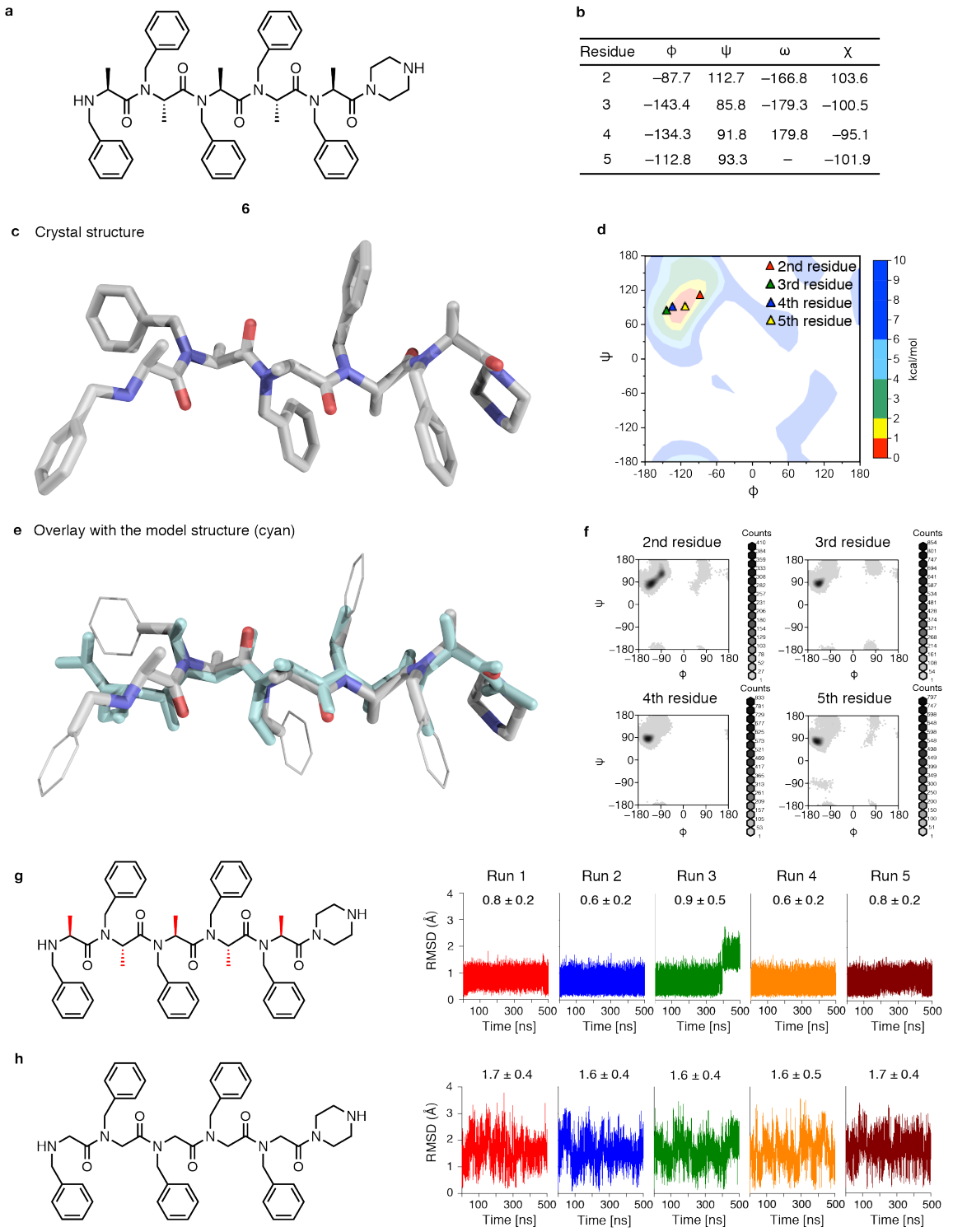


Figure 4 | X-ray crystallographic analysis of oligo-NSA and MD simulations of the structure. (a) Chemical structure of *N*-benzylalanine pentamer **6** that was subjected to the X-ray crystallographic analysis. **(b)** Dihedral angles of oligo-NSA **6** observed in the crystal. **(c)** Crystal structure of oligo-NSA **6**. An ORTEP diagram of the structure is shown in **Figure S9**. **(d)** Dihedral angles observed in the crystal structure were spotted on the Ramachandran-type plot of acetyl-*N*-methylalanine dimethylamide (**Figure 2b**). **(e)** An overlay of the crystal structure and the model structure shown in **Figure 2f** (cyan). **(f)** Distribution of dihedral angles of each residue of oligo-NSA **6** during the last 400 ns (out of 500 ns) × 5 of MD calculations. The calculations were conducted using the crystal structure as the initial conformation. **(g)** RMSD value of the oligo-NSA **6**

during each MD runs. The average and standard deviations are shown above each plot. **(h)** The same MD calculations were conducted for oligo-NSG using the same initial backbone conformation with oligo-NSA **6**. RMSD value of the oligomer from the initial conformation is plotted.

Spectroscopic studies

To evaluate the solution structure of oligo-NSA, NMR studies were conducted. To obtain information about spatial proximity of protons from 2D-NMR spectrum, several heteropentamers that are expected to exhibit dispersed proton peaks on NMR spectrum were designed and synthesized. Among those, NSA heteropentamer **7** satisfied the requirement and all the α-protons of the oligomer and N_α protons on 1H NMR spectrum (**Figure S14**) were successfully assigned unambiguously for this oligomer using COSY (**Figure S15**), 13C NMR (**Figure S16**) and HMBC spectra (**Figure S17**). Although the *N*-substituents are different between oligomer **6** that is used for crystallographic analysis and oligomer **7** that is used for NMR analysis, backbone conformation must be closely resembled with each other, not to say exactly the same, because the extended backbone conformation is dictated without the aid of *N*-substituents as described above. To obtain information about spatial proximities among backbone atoms in an oligomer, NOESY spectrum of the pentamer **7** was measured (**Figure 5a** and **Figure S18** and **S19**). To evaluate backbone conformation, inter-residue NOEs among protons on α, β, and N_α positions were analyzed. As a result, first, all the amide bonds were confirmed to be in trans from inter-residue NOEs between protons on an N_α-carbon and the backbone α-proton of preceding alanine residue (**Figure 5a left**). Restriction of bond rotations about φ angles by pseudo-1,3-allylic strain is expected to result in the intra-residue β and N_α protons to be close with each other while intra-residue α and N_α protons to be away from each other. Existence of strong NOEs between intra-residue β and N_α protons and absence of NOEs between intra-residue α and N_α protons support the restricted rotation about φ angles around the value observed at the crystal structure (**Figure 5a middle**). Similarly, restriction of bond rotations about ψ angles by pseudo-1,3-allylic strain is expected to result in the N_α protons and α proton of preceding alanine residue to be close with each other while N_α protons and β protons of preceding alanine to be away from each other. Existence of strong inter-residue NOEs between N_α protons and α proton of preceding alanine residue support the restricted rotation about ψ angles around the value observed at the crystal structure (**Figure 5a right**). Validity of the interpretation of the observed NOEs was ensured with QM calculations of a model NSA dimer where a φ or ψ angle was systematically rotated and distances between N_α protons and α or β protons on each optimized structure were measured (**Figure S20**). Because β protons of 4th residue exhibited a weak NOE with the N_α protons of 5th residue despite the fact that those protons are more than 3 Å away from each other in the crystal structure (**Figure 5a right, red dotted lines**), some degree of rotation may be allowed for the ψ angle. Together with the absence of NOEs among non-neighboring remote residues, the NMR results overall indicate that the extended conformation of oligo-NSAs seen in the crystal is persistent in aqueous solution.

We also conducted a CD spectroscopic analysis of oligo-NSA to evaluate the effects of the *N*-substituents, concentration, solvent and length of the oligomer on conformation of oligo-NSA. First, CD spectra of oligomer **7**, another heteropentamer **8**, and a homopentamer **5** were measured to

check the sequence dependency of the backbone conformation. All the three oligomers exhibited strong signals and a similar spectral shape that have a maximum around at 195 nm and a minimum around at 225 nm (**Figure 5b**). This result indicates a specific set of backbone dihedral angles is repeated in an oligomer that is independent from *N*-substituent structures. This is consistent with the highly restricted backbone dihedral angles in the QM calculation. The spectrum of homopentamer **5** did not show concentration dependence of the oligomer that eliminates the possibility that the ordered structure is realized by aggregation of the oligomer (**Figure S21a**). The spectral shape did not change in acetonitrile, less polar solvent than water (**Figure S21b**), validating that the conformation is realized by local steric effects on backbone as expected and not dictated by other remote interactions such as hydrophobic interactions or electronic interactions among *N*-substituents. Shorter oligomers, tetramer and trimer, exhibited similar spectra to the pentamers and even dimer exhibited a similar spectral shape although the intensity was little weaker than longer oligomers (**Figure 5c**). This result suggests that an NSA residue is intrinsically restricted in rotations about backbone dihedral angles and oligomers with sequential NSA residues uniformly exhibit a constrained conformation composed of repetition of a similar set of backbone dihedral angles, most probably around (φ, ψ) = (–120°, 90°) that was found to be the lowest energy point in the QM calculation.

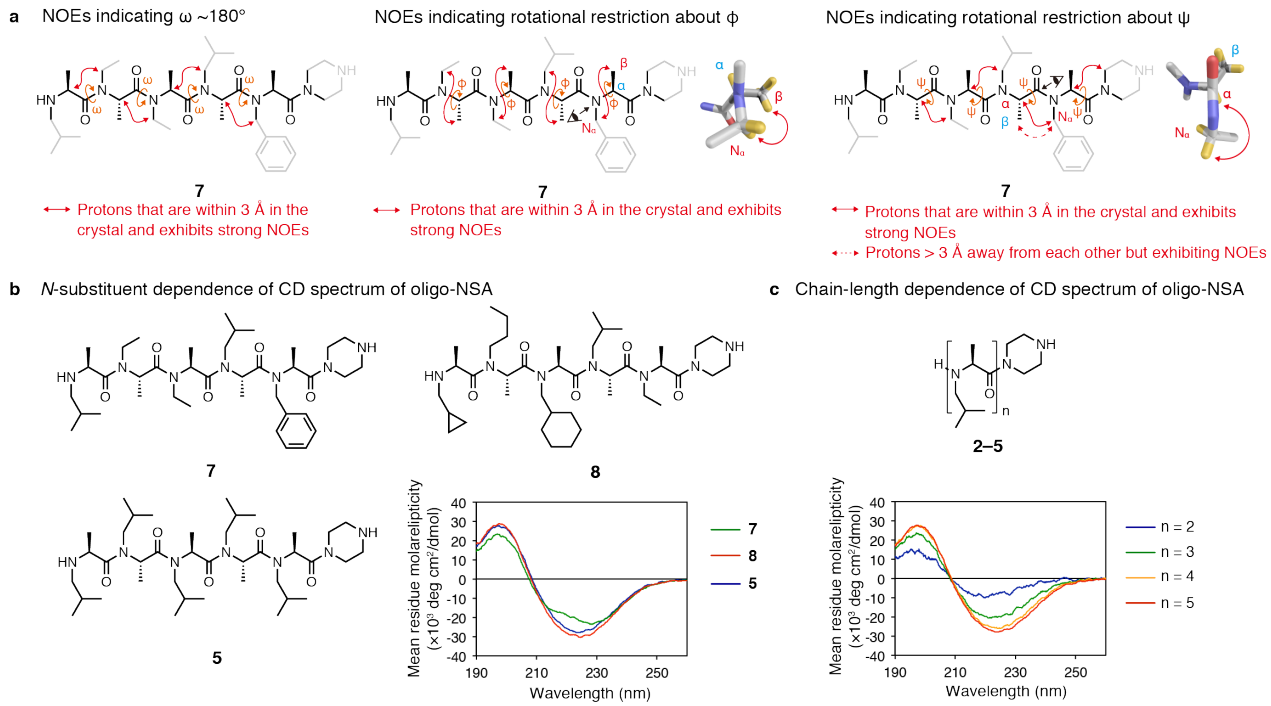


Figure 5 | Conformational studies of oligo-NSAs in aqueous solution. (a) A summary of spatial proximity information obtained from a NOESY measurement. Protons with strong NOEs that suggest rotational restrictions about ω (left), φ (middle), and ψ (right) are indicated with red solid arrows. Protons with a weak NOE related to ψ (right) are indicated with a red dotted arrow. 1H NMR spectrum, COSY spectrum, 13C NMR spectrum, HMBC spectrum and NOESY spectrum are listed on **Figure S14–19**. For the middle and right figures, a part of the crystal structure is shown to visualize the proximities of N_α protons and β protons (middle) or N_α protons and α proton (right). **(b)** Chemical structures and CD spectra of NSA pentamers **7**, **8**, and **5**. **(c)** Chemical structures and CD spectra of oligo-NSA **2**–**5**. The CD spectra were recorded at 25 °C with 100 μM solution of each oligomer in phosphate buffer (pH 7.2). The Y-axis was normalized to molar ellipticity per residue.

Conclusions

As a conclusion, we have established a synthetic scheme of oligo-NSA and conducted a conformational study of oligo-NSA for the first time. The conformational study supported that oligomers with sequential NSA residues form an extended shape in aqueous solution. The extended shape is dictated by backbone steric effects and not dependent on *N*-substituent structures, therefore the oligomer does not collapse into a more compact shape like peptides or other flexible peptidomimetic oligomers.[35–37] This is a stark contrast with NSG-type peptoid that has intrinsically flexible backbone structure and requires elaborate *N*-substituents to achieve constrained conformation. The submonomeric synthetic method of oligo-NSAs allows facile introduction of a diverse set of functional groups as *N*-substituents using readily available aldehydes. These features make oligo-NSA as an ideal scaffold for displaying functional groups into a well-defined three-dimensional space in aqueous solution that realizes efficient biomolecular recognition and catalytic activity.

References

1. Farmer, P. & Ariëns, E. Speculations on the design of nonpeptidic peptidomimetics. *Trends Pharmacol. Sci.* **4,** 362–365 (1982).
2. Horwell, D. C. Use of the chemical structure of peptides as the starting point to design nonpeptide agonists and antagonists at peptide receptors: Examples with cholecystokinin and tachykinins. *Bioorg. Med. Chem.* **4,** 1573–1576 (1996).
3. Simon, R. J. *et al.* Peptoids: a modular approach to drug discovery. *Proc. Natl. Acad. Sci. U. S. A.* **89,** 9367–9371 (1992).
4. Miller, S. M. *et al.* Comparison of the proteolytic susceptibilities of homologous L-amino acid, D-amino acid, and N-substituted glycine peptide and peptoid oligomers. *Drug Dev. Res.* **35,** 20–32 (1995).
5. Yu, P., Liu, B. & Kodadek, T. A high-throughput assay for assessing the cell permeability of combinatorial libraries. *Nat. Biotechnol.* **23,** 746–751 (2005).
6. Zuckermann, R. N., Kerr, J. M., Kent, S. B. H. & Moos, W. H. Efficient Method for the Preparation of Peptoids [Oligo(N-substituted glycines)] by submonomer solid-phase synthesis. *J. Am. Chem. Soc.* **114,** 10646–10647 (1992).
7. Pels, K. & Kodadek, T. Solid-Phase Synthesis of Diverse Peptide Tertiary Amides By Reductive Amination. *ACS Comb. Sci.* **17,** 152–155 (2015).
8. Zuckermann, R. N. *et al.* Discovery of Nanomolar Ligands for 7-Transmembrane G-Protein-Coupled Receptors from a Diverse N-(Substituted)glycine Peptoid Library. *J. Med. Chem.* **37,** 2678–2685 (1994).
9. Wu, C. W., Sanborn, T. J., Huang, K., Zuckermann, R. N. & Barron, A. E. Peptoid oligomers with α-chiral, aromatic side chains: Sequence requirements for the formation of stable peptoid helices. *J. Am. Chem. Soc.* **123,** 6778–6784 (2001).
10. Shah, N. H. *et al.* Oligo(N-aryl glycines): A new twist on structured peptoids. *J. Am. Chem. Soc.* **130,** 16622–16632 (2008).
11. Stringer, J. R., Crapster, J. A., Guzei, I. A. & Blackwell, H. E. Extraordinarily robust polyproline type i peptoid helices generated via the incorporation of α-chiral aromatic N -1-naphthylethyl side chains. *J. Am. Chem. Soc.* **133,** 15559–15567 (2011).
12. Gorske, B. C., Stringer, J. R., Bastian, B. L., Fowler, S. A. & Blackwell, H. E. New strategies for the design of folded peptoids revealed by a survey of noncovalent interactions in model systems. *J. Am. Chem. Soc.* **131,** 16555–16567 (2009).
13. Gorske, B. C., Mumford, E. M. & Conry, R. R. Tandem Incorporation of Enantiomeric Residues Engenders Discrete Peptoid Structures. *Org. Lett.* **18,** 2780–2783 (2016).
14. Gorske, B. C., Mumford, E. M., Gerrity, C. G. & Ko, I. A Peptoid Square Helix via Synergistic Control of Backbone Dihedral Angles. *J. Am. Chem. Soc.* **139,** 8070–8073 (2017).
15. Sanborn, T. J., Wu, C. W., Zuckermann, R. N. & Barron, A. E. Extreme stability of helices

formed by water-soluble poly-N-substituted glycines (polypeptoids) with α-chiral side chains. *Biopolymers* **63,** 12–20 (2002).

16. Shin, S. B. Y. & Kirshenbaum, K. Conformational rearrangements by water-soluble peptoid foldamers. *Org. Lett.* **9,** 5003–5006 (2007).
17. Fuller, A. A., Yurash, B. A., Schaumann, E. N. & Seidl, F. J. Self-association of water-soluble peptoids comprising (S)-N-1- (naphthylethyl)glycine residues. *Org. Lett.* **15,** 5118–5121 (2013).
18. Darapaneni, C. M., Kaniraj, P. J. & Maayan, G. Water soluble hydrophobic peptoids via a minor backbone modification. *Org. Biomol. Chem.* **16,** 1480–1488 (2018).
19. Gao, Y. & Kodadek, T. Synthesis and screening of stereochemically diverse combinatorial libraries of peptide tertiary amides. *Chem. Biol.* **20,** 360–369 (2013).
20. Kodadek, T. & McEnaney, P. J. Towards vast libraries of scaffold-diverse, conformationally constrained oligomers. *Chem. Commun.* **52,** 6038–6059 (2016).
21. Hoffmann, R. W. Allylic 1,3-Strain as a Controlling Factor in Stereoselective Transformations. *Chem. Rev.* **89,** 1841–1860 (1989).
22. Hoffmann, R. W. Flexible Molecules with Defined Shape-Conformational Design. *Angew. Chem. Int. Ed.* **31,** 1124–1134 (1992).
23. Moehle, K. & Hofmann, H. J. Peptides and Peptoids-A Quantum Chemical Structure Comparison. *Biopolymers* **38,** 781–790 (1996).
24. Butterfoss, G. L., Renfrew, P. D., Kuhlman, B., Kirshenbaum, K. & Bonneau, R. A preliminary survey of the peptoid folding landscape. *J. Am. Chem. Soc.* **131,** 16798–16807 (2009).
25. Zuckermann, R. N. *et al.* Synthesis of N-substituted oligomers. *US Pat. Number US 5831005* (1998).
26. Virgilio, A. A. & Ellman, J. A. Simultaneous Solid-Phase Synthesis of β-Turn Mimetics Incorporating Side-Chain Functionality. *J. Am. Chem. Soc.* **116,** 11580–11581 (1994).
27. Thern, B., Rudolph, J. & Jung, G. Total Synthesis of the Nematicidal Cyclododecapeptide OmphalotinA by Using Racemization-Free Triphosgene-Mediated Couplings in the Solid Phase. *Angew. Chem. Int. Ed.* **41,** 2307–2309 (2002).
28. Chatterjee, J., Laufer, B. & Kessler, H. Synthesis of N-methylated cyclic peptides. *Nat. Protoc.* **7,** 432–444 (2012).
29. Morimoto, J. & Kodadek, T. Synthesis of a large library of macrocyclic peptides containing multiple and diverse N-alkylated residues. *Mol. Biosyst.* **11,** 2770–2779 (2015).
30. Kaminker, R. *et al.* Tuning conformation and properties of peptidomimetic backbones through dual: N / Cα-substitution. *Chem. Commun.* **54,** 5237–5240 (2018).
31. Kaminker, R. *et al.* Tuning of protease resistance in oligopeptides through N-alkylation. *Chem. Commun.* **54,** 9631–9634 (2018).

32. Fernández-Llamazares, A. I. *et al.* N-Triethylene glycol (N-TEG) as a surrogate for the N-methyl group: application to Sansalvamide A peptide analogs. *Chem. Commun.* **49,** 6430–6432 (2013).
33. Fernández-Llamazares, A. I. *et al.* The backbone N-(4-azidobutyl) linker for the preparation of peptide chimera. *Org. Lett.* **15,** 4572–4575 (2013).
34. Sui, Q., Borchardt, D. & Rabenstein, D. L. Kinetics and equilibria of cis/trans isomerization of backbone amide bonds in peptoids. *J. Am. Chem. Soc.* **129,** 12042–12048 (2007).
35. Lau, K. F. & Dill, K. A. A Lattice Statistical Mechanics Model of the Conformational and Sequence Spaces of Proteins. *Macromolecules* **22,** 3986–3997 (1989).
36. Wiley, R. A. & Rich, D. H. Peptidomimetics Derived from Natural Products. *Med. Res. Rev.* **13,** 327–384 (1993).
37. Bradley, E. K. *et al.* NMR structural characterization of oligo-N-substituted glycine lead compound for a combinatorial library. *Mol. Divers.* **3,** 1–15 (1997).

Acknowledgements

J.M. acknowledges financial support from JSPS (JP17K13265). S.S. acknowledges financial support from CREST (JPMJCR13L4), Japan Science and Technology Agency. D.K. acknowledges financial support from JSPS (JP17K18113) and AMED (JP18fm0208022h). K.T. acknowledges financial support from JSPS (JP16H02420 and JM17H06109) and Japan Agency for Medical Research and development, AMED (JP18arn0101094j). Y.F. was supported from the Graduate Program for Leaders in Life Innovation from MEXT, Japan. The computations were performed using Research Center for Computational Science, Okazaki, Japan. We thank Dr. H. Sato Mr. S. Suginome, Dr. S. Kusumoto, Dr. M. Akiyama, and Dr. Okitsu at the University of Tokyo and Dr. H. Sato at Rigaku for the support on X-ray crystallographic analysis. We thank Drs. Y. Goto and H. Suga at the University of Tokyo for the use of a CD spectrometer. The X-ray diffraction data of the crystal was collected using an instrument at Advanced Characterization Nanotechnology Platform of the University of Tokyo, supported by "Nanotechnology Platform" of the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

Author contributions

J.M. and S.S. conceived and directed the project. J.M. and Y.F. contributed equally to the execution of this work. Y.F. and J.M. synthesized compounds. Y.F. performed the X-ray crystallographic study. Y. F. and T.W. conducted solution phase structural studies of oligo-NSA. J.M. and D.K. designed and conducted quantum mechanical calculations. D.K. and K.T. designed and conducted molecular dynamics simulations. All authors contributed to analysis of the obtained results and writing the manuscript.

Competing interests

The authors declare no competing financial interests.

Materials & Correspondence

Supplementary information containing additional figures and tables and compoud characterization data is available as a separate file. Correspondence should be addressed to J.M. (jmorimoto@chembio.t.u-tokyo.ac.jp) and S.S. (ssando@chembio.t.u-tokyo.ac.jp).

Methods

General remarks. Chemicals and solvents used in this study were purchased from commercial suppliers and used without further purification. Synthesis under microwave conditions was performed on Initiator+ (Biotage). Preparative HPLC was performed on a Prominence HPLC system (Shimazu) with a 5C18-MS-II column (Nacalai tesque, 10 mm I.D.×150 mm, 34355-91). Analytical HPLC was performed on a Prominence HPLC system with a 5C18-AR-II column (Nacalai tesque, 4.6 mm I.D.×150 mm, 38144-31). HRMS data was obtained using micrOTOF II (Bruker Daltonics). MALDI-TOF MS analysis was performed on autoflex speed (Bruker Daltonics) using super-DHB (Santa Cruz Biotechnology) as matrix. All quantum calculations were carried out with the Gaussian16 package[38] and all molecular dynamics simulations were carried out with the Gromacs 2018 package and the CHARMM36m force field.

Generation of Ramachandran-type energy diagrams of NSG and NSA. Ramachandran-type energy diagrams[39] of acetyl-*N*-methylglycine dimethylamide and acetyl-*N*-methylalanine dimethylamide were generated by combinatorialy fixing φ and ψ at every 15° from –180° to 180°. Each conformer was optimized at the B3LYP/6-31G* level. The ω angle was fixed to 180° through the calculation. Ramachandran-type energy landscapes of acetyl-*N*-ethylglycine dimethylamide was generated by combinatorialy fixing φ and ψ at every 15° from –180° to 180°. Each conformer was optimized at the B3LYP/6-31G* level. The ω angle was fixed to 180° through the calculation. Calculations were performed for two molecule with a fixed χ angle of either 100° or –100°.

χ scan of acetyl-*N*-ethylalanine dimethylamide. χ scan was performed by optimizing a conformer with χ angle of –180° to 180° with 10° increment. The calculation started with an initial conformation of (χ, φ, ψ) = (0°, –120°, 90°).

Computationally optimized conformations of a model NSA pentamer. Conformations of acetyl-*N*-ethylalanine pentamer dimethylamide with dihedral angles (χ, φ, ψ, ω) of either (100°, –105°, 105°, 180°) or (–100°, –120°, 90°, 180°) were optimized by quantum mechanical calculations. The two sets of the angles are the ones that were calculated to be the lowest energy angles in the Ramachandran-type diagrams of acetyl-*N*-ethylalanine dimethylamide. Each conformer was optimized at the B3LYP/6-31G* level using a self-consistent reaction field (SCRF) model with water as the solvent. Single point energy of the optimized conformer was also calculated at the B3LYP/6-31G* level using a SCRF model with water as the solvent.

Synthesis of NSA-containing peptide using bis(trichloromethyl)carbonate (BTC) as a coupling reagent. Rink amide ChemMatrix resin was swelled with minimal volume of *N*,*N*-dimethylformamide (DMF) in a syringe for 30 min. DMF was filtered off and the resin was

treated with DMF solution of Fmoc-Trp(Boc)-OH (4 equiv., 0.2 M), 1-[(1-(cyano-2-ethoxy-2-oxoethylideneaminooxy) dimethylaminomorpholino)] uronium hexafluorophosphate (COMU) (4 equiv., 0.2 M) and DIPEA (*N*,*N*-diisopropylethylamine) (8 equiv., 0.4 M) with continuous shaking for 2 h. After removing solution, the resin was washed with DMF three times. The resin was treated with 20% piperidine/DMF (3 min and 12 min) and then washed with DMF and tetrahidrofuran (THF) three times each. The resin was incubated with 27 equiv. of DIPEA in anhydrous THF for 15 min. A solution of BTC (4 equiv., 0.1 M) and (*R*)-2-bromopropionic acid or (*S*)-2-bromopropionic acid (0.1 M, 12 equiv.) in anhydrous THF was cooled at –20 °C for 15 min and mixed with 36 equiv. of 2,4,6-Trimethylpyridine. The solution was immediately applied to the resin and the reaction mixture was shaken for 2 h. After the reaction, the resin was washed with THF and DMF three times each. The resin was incubated with 2 M (80 equiv.) of an isobutylamine solution in anhydrous DMF for 18 h at 60 °C. The solution was filtered off and the resin was washed with DMF and dichloromethane (DCM) three times each. The peptides were cleaved by treating the resin with 95/5 TFA (trifluoroacetic acid)/H_2O for 2 h. The solution was transferred to a recovery flask and TFA solution was removed under reduced pressure. The crude product was dissolved in 10/90 acetonitrile/water and analyzed by a reversed phase column on HPLC.

Investigation about coupling conditions and evaluation of racemization during synthesis. Rink amide ChemMatrix resin was swelled with minimal volume of DMF in a syringe for 30 min. DMF was filtered off and the resin was treated with DMF solution of Fmoc-Trp(Boc)-OH (4 equiv., 0.2 M), COMU (4 equiv., 0.2 M) and DIPEA (8 equiv., 0.4 M) with continuous shaking for 40 min. After removing solution, the resin was washed with DMF three times. The resin was treated with 20% piperidine/DMF (3 min and 12 min) and then washed with DMF three times. After deprotection, the resin was treated with DMF solution of Fmoc-Ala-OH (4 equiv., 0.2 M), COMU (4 equiv., 0.2 M) and DIPEA (8 equiv., 0.4 M). After 40 min, the resin was washed with DMF and Fmoc group was removed by treatment with 20% piperidine/DMF. The resin was incubated with 1 M (20 equiv.) of an isobutylaldehyde solution in anhydrous DMF. Aldehyde solution was filtered off and the resin was quickly washed with DMF and DCM. A freshly prepared suspension of $NaBH_4$ (10 equiv.) in 75/25 DCM/MeOH was added to the resin and shaken for 30 min. The cap of the syringe was occasionally detached for degassing to prevent inner pressure is increased too much. After reduction, the resin was washed with MeOH five times then with DCM, DMF and dioxane three times each. When using COMU or 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate (HATU) as a coupling reagent, a solution of Fmoc-L-Ala-OH (4 equiv., 0.2 M), coupling reagent (4 equiv., 0.2 M) and DIPEA (8 equiv., 0.4 M) in anhydrous DMF was added to the resin and shaken under the condition shown in **Figure S3**. After the reaction, the resin was washed with DMF three times. When BTC was used as the coupling reagent, the resin was washed with THF three times and incubated with 8 equiv. of DIPEA in anhydrous THF for 15 min. A solution of

Fmoc-L-Ala-OH (3.5 equiv., 0.1 M) and BTC (1.17 equiv., 0.1 M) in anhydrous THF was cooled at –20 °C for 15 min and mixed with 10 equiv. of 2,4,6-trimethylpyridine. The solution was immediately applied to the resin and the reaction mixture was shaken under the condition shown in **Figure S3**. After the reaction, the resin was washed with THF and DMF three times each. When using 1-ethoxycarbonyl-2-ethoxy-1,2-dihydroquinoline (EEDQ), the resin was washed with dioxane three times and a solution of Fmoc-L-Ala-OH or Fmoc-D-Ala-OH (4 equiv., 0.2 M) and EEDQ (4 equiv., 0.2 M) in dioxane was added to the resin and shaken under the condition shown in **Figure S3**. After the reaction, the resin was washed with dioxane and DMF three times each. After the coupling reaction to *N*-substituted alanine terminus with each coupling reagent, Fmoc protecting group was removed with 20% piperidine/DMF. The peptides were cleaved by treating the resin with 95/5 TFA/H_2O for 2 h. The solution was transferred to a flask and TFA solution was removed under reduced pressure. The crude product was dissolved in acetonitrile and water and analyzed by a reversed phase column on HPLC.

General procedure for oligo-NSA synthesis. Trityl resin was swelled with minimal volume of THF in a flask for 10 min. THF solution of piperazine (4 equiv., 0.2 M) and piperidine (16 equiv., 0.8 M) was added to the resin. Piperidine was added in order to reduce the reactive points on the resin. After stirring for 2 h, resin was moved to syringe and washed with THF and DCM three times each. The resin was treated with 85/10/5 DCM/MeOH/DIPEA solution for 15 min and washed with DCM and DMF three times each. The resin was treated with DMF solution of Fmoc-Ala-OH (4 equiv., 0.2 M), COMU (4 equiv., 0.2 M) and DIPEA (8 equiv., 0.4 M) with continuous shaking for 2 h. After removing solution, the resin was washed with DMF and DCM three times each and dried. A small part of the resin was treated with 2% 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU)/DMF solution and then the load amount of Fmoc-alanine was determined from the UV absorbance of dibenzofulvene at 304 nm as described in reference[40]. In the following procedures, the amount of reagents was calculated based on the determined load amount. After Fmoc quantification, the resin was swelled with minimal volume of DMF for 30 min. The resin was treated with 20% piperidine/DMF (3 min and 12 min) to remove Fmoc protecting group and washed with DMF three times. The resin was incubated with 1 M (20 equiv.) of an aldehyde solution in anhydrous DMF. Aldehyde solution was filtered off and the resin was quickly washed with DMF and DCM. A freshly prepared suspension of $NaBH_4$ (10 equiv.) in 75/25 DCM/MeOH was added to the resin and shaken for 30 min. The cap of the syringe was occasionally detached for degassing to prevent inner pressure is increased too much. After reduction, the resin was washed with MeOH five times then with DCM, DMF and dioxane three times each. Fmoc-Ala-OH (4 equiv.) and EEDQ (4 equiv.) were dissolved in dioxane to prepare 0.2 M solution and the mixture was shaken. After 30 min, the mixture was added to the resin and shaken for 3 h at 60 °C. This coupling reaction was repeated once more. After double coupling, Fmoc protecting group was removed with 20% piperidine/DMF and substituents were introduced by reductive

amination. The coupling, deprotection and reductive amination procedures were repeated to afford objective oligomers on resin. The oligomers were cleaved by treating the resin with 30% 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP)/DCM for 20min five times. The solution was transferred to a glass vial and HFIP solution was removed under reduced pressure. The crude product was dissolved in acetonitrile and water and purified by a reversed phase column on HPLC. The yields were calculated by comparing molecular weight of each obtained compound as TFA salt and determined load amount of the first Fmoc-alanine. The crude and purified products were analyzed on a reversed phase column by HPLC and ESI-TOF MS.

***N*-Isobutylalanine piperazineamide (1).** 42 mg of trityl resin (1.96 mmol/g, 83 μmol) was used for synthesis. Fmoc-alanine and isobutylaldehyde were used as submonomers. After loading first Fmoc-alanine, load amount was determined as 0.48 mmol/g. Yield was 7.2 mg (80%). ^{1}H NMR (CD_3CN, 400 MHz): δ 1.00–1.04 (q, 6H, J = 6.9 Hz), 1.55 (d, 3H, J = 6.9), 1.97–2.10 (m, 1H), 2.74–2.82 (dd, 1H, J = 7.3, 12.4), 2.74–2.99 (dd, 1H, J = 6.9, 12.4), 3.33-3.43 (m, 4H), 3.83–3.97 (m, 4H), 4.48 (q, 1H, J = 6.9 Hz). ^{13}C NMR (CD_3CN, 400 MHz): δ 14.7, 19.1, 19.2, 25.9, 39.1, 42.0, 42.8, 42.8, 53.9, 54.4, 168.3. The ^{1}H NMR is shown in **Figure S7** and ^{13}C NMR spectra is shown in **Figure S8.** HRMS (ESI-TOF MS) *m/z*: $[M + H]^+$ Calcd for $C_{11}H_{24}N_3O^+$ 214.1914; Found 214.1929.

***N*-Isobutylalanine dimer piperazineamide (2).** 42 mg of trityl resin (1.96 mmol/g, 83 μmol) was used for synthesis. Fmoc-alanine and isobutylaldehyde were used as submonomers. After loading first Fmoc-alanine, load amount was determined as 0.48 mmol/g. Yield was 8.2 mg (71%). HRMS (ESI-TOF MS) *m/z*: $[M + H]^+$ Calcd for $C_{18}H_{37}N_4O_2^+$ 341.2911; Found 341.2889.

***N*-Isobutylalanine trimer piperazineamide (3).** 41 mg of trityl resin (1.96 mmol/g, 80 μmol) was used for synthesis. Fmoc-alanine and isobutylaldehyde were used as submonomers. After loading first Fmoc-alanine, load amount was determined as 0.33 mmol/g. Yield was 5.1 mg (54%). HRMS (ESI-TOF MS) *m/z*: $[M + H]^+$ Calcd for $C_{25}H_{50}N_5O_3^+$ 468.3908; Found 468.3916.

***N*-Isobutylalanine tetramer piperazineamide (4).** 39 mg of trityl resin (1.96 mmol/g, 76 μmol) was used for synthesis. Fmoc-alanine and isobutylaldehyde were used as submonomers. After loading first Fmoc-alanine, load amount was determined as 0.46 mmol/g. Yield was 8.1 mg (55%). HRMS (ESI-TOF MS) *m/z*: $[M + H]^+$ Calcd for $C_{32}H_{63}N_6O_4^+$ 595.4905; Found 595.4905.

***N*-Isobutylalanine pentamer piperazineamide (5).** 42 mg of trityl resin (1.96 mmol/g, 83 μmol) was used for synthesis. Fmoc-alanine and isobutylaldehyde were used as submonomers. After loading first Fmoc-alanine, load amount was determined as 0.48 mmol/g. Yield was 9.2 mg (47%). HRMS (ESI-TOF MS) *m/z*: $[M + H]^+$ Calcd for $C_{39}H_{76}N_7O_5^+$ 722.5902; Found 722.5890.

***N*-Benzylalanine pentamer piperazineamide (6).** 38 mg of trityl resin (1.96 mmol/g, 75 μmol) was used for synthesis. Fmoc-alanine and benzaldehyde were used as submonomers. After loading first Fmoc-alanine, load amount was determined as 0.56 mmol/g. Yield was 6.4 mg (27%). HRMS (ESI-TOF MS) *m/z*: $[M + H]^+$ Calcd for $C_{54}H_{66}N_7O_5^+$ 892.5120; Found 892.5132.

NSA heteropentamer (7). 32 mg of trityl resin (1.96 mmol/g, 63 μmol) was used for synthesis. Fmoc-alanine, benzaldehyde, isobutylaldehyde and acetaldehyde were used as submonomers. After loading first Fmoc-alanine, load amount was determined as 0.51 mmol/g. Yield was 6.2 mg (41%). HRMS (ESI-TOF MS) *m/z*: $[M + H]^+$ Calcd for $C_{38}H_{66}N_7O_5^+$ 700.5120; Found 700.5120.

NSA heteropentamer (8). 40 mg of trityl resin (1.96 mmol/g, 78 μmol) was used for synthesis. Fmoc-alanine, acetaldehyde, isobutylaldehyde, cyclohexanecarboxaldehyde, n-butylaldehyde and cyclopropanecarboxaldehyde were used as submonomers. After loading first Fmoc-alanine, load amount was determined as 0.38 mmol/g. Yield was 4.9 mg (35%). HRMS (ESI-TOF MS) *m/z*: $[M + H]^+$ Calcd for $C_{40}H_{74}N_7O_5^+$ 732.5746; Found 732.5742.

Crystallization. A lyophilized product of *N*-benzylalanine pentamer piperazineamide (**6**) was dissolved in a mixture of hexane and dichloromethane. The solution was left in a glass vial capped with a plastic cap with small holes to let solvent slowly evaporate until crystals appeared in solution.

X-ray crystallography. A single crystal was mounted with mineral oil on a loop-type mount and set on VariMax Dual (Rigaku). The X-ray diffraction data was measured at –180 °C using Mo *Kα* radiation (λ = 0.7107 Å). Data was processed using the CrystalClear software (Rigaku). The structure was solved by a direct method using SHELXT[41] and refined using SHELXL. The non-hydrogen atoms were refined anisotropically. Hydrogen atoms were placed on ideal positions.

Molecular dyanamics simulations. MD simulations of the oligo-NSA, the oligo-NSG and the oligomer of alternate NSA and NSG residues were performed using GROMACS 2018.1[42] with the CHARMM36m force field and the CMAP correction[43] and the TIP3P water model.[44] The force field for the peptoid molecules was derived from the CGenFF.[45] The initial structure of each simulation was derived from the crystal structure of NSA and was solvated with TIP3P water in a rectangular box such that the minimum distance to the edge of the box was 15 Å under periodic boundary conditions through the CHARMM-GUI.[45] Na and Cl ions were added to imitate a salt solution of concentration 0.14 M. The system was energy-minimized for 10,000 steps and equilibrated with the NVT ensemble (298 K) for 1 ns. Further simulations were performed with the NPT ensemble at 298 K for 500 ns. For each system, the simulation was repeated 5 times with different initial velocities (i.e. 2.5

μs in total for each peptoid). The time step was set to 2 fs throughout the simulations. A cutoff distance of 12 Å was used for Coulomb and van der Waals interactions. Long-range electrostatic interactions were evaluated by means of the particle mesh Ewald method.[46] Covalent bonds involving hydrogen atoms were constrained by the LINCS algorithm.[47] A snapshot was saved every 10 ps. For the analysis of each trajectory, we employed the last 400 ns.

NMR spectroscopic studies. NMR spectra of NSA pentamer **7** were recorded at 5 mM in D_2O on a JEOL ECS-400. 1H NMR and 13C NMR spectrum are shown in **Figure S14** and **Figure S16.** COSY spectrum was recorded with relaxation delay of 1.5 s and receiver gain of 42 (**Figure S15**). HMBC spectrum was recorded with x points of 2048, y points of 2048, relaxation delay of 1.5 s and receiver gain of 90 (**Figure S17**). NOESY spectrum was recorded with relaxation delay of 1.5 s, mixing time of 0.8 s and receiver gain of 42 (**Figure S18** and **S19**). Chemical shifts of 1H NMR, HMBC, COSY and NOESY spectrum are reported in p.p.m relative to solvent peaks as internal standards (δH, H_2O 4.79 ppm). Assignment of 1H NMR was assisted by COSY and HMBC spectrum. Sequential assignment of main chain alpha protons was completed from their intra-residual and inter-residual cross peaks with carbonyl carbon in HMBC spectrum. Sequential assignment of alpha protons on *N*-substituents was completed from their inter-residual cross peaks with carbonyl carbon in HMBC spectrum. All assignment was shown in **Figure S14**. The obtained cross peaks were interpreted based on QM calculations of a model NSA dimer. More specifically, acetyl-*N*-ethylalanine dimer dimethylamide with dihedral angles (χ, φ, ψ, ω) of (–100°, –120°, 90°, 180°) was systematically rotated at the φ angle of C-terminal residue or ψ angle of *N*-terminal residue with 15° increment at a time and optimized at the B3LYP/6-31G* level using a SCRF model with water as the solvent. The distances between α or β protons and N_α protons on each conformer were listed as tables on **Figure S20**.

Circular dichroism studies. Solution of oligo-NSA was prepared by dissolving lyophilized compounds in water or acetonitrile followed by diluting to the desired concentration solution in 10 mM phosphate buffer or acetonitrile. CD spectra were acquired at 25 °C with a CD spectrometer (JASCO, J-1500) using 1 mm path length quartz cell (JASCO, 209J). Data pitch was set to 0.1 mm. The scanning speed was set to 100 nm/min and spectra were averaged from three scans. Spectral baseline was recorded using 10 mM phosphate buffer or acetonitrile. All data points were baseline subtracted, converted to a uniform scale of molar ellipticity per residue and plotted.

References

38. Frisch, M. J. *et al.* Gaussian 16, Revision B.01.
39. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7,** 95–99 (1963).

40. Gude, M., Ryf, J. & White, P. D. An accurate method for the quantitation of Fmoc-derivatized solid phase supports. *Lett. Pept. Sci.* **9,** 203–206 (2002).
41. Sheldrick, G. M. SHELXT - Integrated space-group and crystal-structure determination. *Acta Crystallogr. Sect. A Found. Crystallogr.* **71,** 3–8 (2015).
42. Abraham, M. J. *et al.* Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2,** 19–25 (2015).
43. Huang, J. *et al.* CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14,** 71–73 (2016).
44. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79,** 926–935 (1983).
45. Brooks, B. R. *et al.* CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **30,** 1545–1614 (2009).
46. Brooks, B. R. *et al.* CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **30,** 1545–1614 (2009).
47. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98,** 10089–10092 (1993).
48. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for molecular simulations. *J. Comput. Chem.* **18,** 1463–1472 (1997).