Visualized and Quantitative Conformational Analysis of Peptidomimetics

Hajime Takashima,* Atsushi Yoshimori, Eiji Honda, Tomonori Taguri, Jun Ozawa, Masaji Kasai, Satoshi Shuto, and Dai Takehara

ABSTRACT: Protein–protein interactions (PPIs) are fundamentally important and challenging drug targets. Peptidomimetic molecules of various types have been developed to modulate PPIs. A particularly promising drug discovery strategy, structural peptidomimetics, was designed based on special mimicking of side-chain $\alpha$-helix $C_{\alpha} - C_\beta$ bonds. It is simple and versatile. Nevertheless, no quantitative method has been established to evaluate its similarity to a target peptide motif. We developed two methods that enable visual, comprehensive, and quantitative analysis of peptidomimetics: peptide conformation distribution (PCD) plot and peptidomimetic analysis (PMA) map. These methods specifically examine multiple side-chain $C_{\alpha} - C_\beta$ bonds of a peptide fragment motif and their corresponding bonds (pseudo-$C_{\alpha} - C_\beta$ bonds) in a mimetic molecule instead of $\phi$ and $\psi$ angles of a single amino acid in the traditional Ramachandran plot. The PCD plot is an alignment-free method, whereas the PMA map is an alignment-based method providing distinctive and complementary analysis. Results obtained from analysis using these two methods indicate our multifaceted $\alpha$-helix mimetic scaffold as an excellent peptidomimetic that can precisely mimic the spatial positioning of side-chain functional groups of $\alpha$-helix. These methods are useful for visualized and quantified evaluation of peptidomimetics and for the rational design of new mimetic scaffolds.

INTRODUCTION

Protein–protein interactions (PPIs) are involved in many regulatory pathways such as signal transduction, receptor–ligand interaction, cell metabolism, and transport across membranes. Often, PPIs are mediated by a “hot spot” amino acid side-chain functionality organized on secondary structures. Modulation and inhibition of PPIs by peptidomimetics is a promising strategy for drug discovery. The $\alpha$-helix, which is the most common secondary structure on the surface of proteins, frequently mediates interactions of proteins with binding partners. In fact, 62% of PPIs in the Protein Data Bank (PDB) have an $\alpha$-helix at the interface, underscoring the importance of the secondary structure for protein–protein recognition. The mimicry of an $\alpha$-helix structure is of considerable interest for drug discovery for PPI modulators.

Efforts to mimic an $\alpha$-helix structure have led to several synthetic strategies such as helix stabilization, helical foldamers, side-chain mimetics, and pharmacophore mimetics on a helical surface. Grossmann et al. introduced four distinctive classifications of peptidomimetics including $\alpha$-helix mimetics: minor modified peptides/stapled peptides (Class A); major modified peptides/foldamers (Class B); structural mimetics (Class C); and mechanistic mimetics/pharmacophore mimetics (Class D). Classes C and D are completely nonpeptide small molecules. Among them, structural mimetics is a definite and versatile drug discovery strategy for providing fine $\alpha$-helix mimetics because the scaffold of a mimetic molecule completely replaces the helical backbone. Key residues are also involved in the mimetics. The basic concept of Class C structural mimetics is that the side-chain $C_{\alpha} - C_\beta$ bond of a peptide fragment is projected to the bond of the side chain in structural mimetics. For this study, we defined the bond mimicking the side-chain $C_{\alpha} - C_\beta$ bond of a peptide fragment as the “pseudo-$C_{\alpha} - C_\beta$ bond” (Figure 1a, red stick). Arora et al. reported that preferred side-chain rotamers of a peptide contribute directly to specificity in protein complex formation, indicating the importance of the $C_{\alpha} - C_\beta$ bond mimetics strategy. In contrast to Class C, the side-chain functional group constituting the pharmacophore of a peptide is involved as pharmacophore mimetics in Class D. The pharmacophore mimetics is connected to the mimetics core...
mimetics. Similarly, the mimic various conformations of peptidomimetics. In these mimetics, single-facial helix mimetics is the for molecular recognition. Similarly, mimetic compounds are designed to inhibit binding of an intrinsically disordered protein to a receptor protein mSin3 by mimicking the NRSF/REST. Compound 5a, designated as mS-11, which has the octahydropyrazino[2,1-c]-1,2,4-oxadiazine scaffold 11. The scaffold can mimic all four side chains involved in a one-turn helix. Compound 12a is designed to inhibit binding of an intrinsically disordered protein NRSF/REST to a receptor protein mSin3 by mimicking the one-turn α-helix motif, Leu46-Ile47-Met48-Leu49 (LIML), of the NRSF/REST. Compound 12a is confirmed to interact with mSin3, inhibiting mSin3-NRSF/REST binding by NMR analysis (PDB ID: SY95). It also ameliorates social interaction deficits in a prenatal valproic acid-induced autism mouse model. Detailed in silico analysis by enhanced molecular dynamics simulation revealed that this scaffold induced PPI inhibition by long-range molecular orientation ordering followed by short-range interactions.

Although many structural peptidomimetics have been developed, no analytical or validation method has been established to demonstrate quantitatively how similarly these molecules can mimic target peptide structures. This report describes two novel methods for analyzing structural peptidomimetics. They enable quantitative and visual understanding of the structural similarity of peptidomimetics to their target peptides based on comparison of side-chain Cα−Cβ bonds and pseudo-Cα−Cβ bonds.

## RESULTS AND DISCUSSION

**PCD Plot: Alignment-Free Method Based on Cα−Cβ Bonds.** For the analysis of protein structures such as secondary structures, Ramachandran plots based on ϕ and ψ angles have often been used as a well-known powerful method. Comprehensive analysis of ϕ and ψ angle distribution in the Ramachandran plot has been reported for classification by secondary structure types and amino acid types using non-redundant protein structure sets. The ϕ and ψ angle analysis is effective for analyzing peptide backbone structures, but it cannot be applied to structural peptidomimetics. The main chain is replaced completely by a small-molecule scaffold in structural peptidomimetics. Therefore, ϕ and ψ angles cannot be defined. This point of inadequacy might make it difficult to evaluate the similarity of structural peptidomimetics qualitatively with its target peptide motif.

Using another structural analysis approach, Garland et al. demonstrated the utility of Cα−Cβ bond vectors for the classification and analysis of β-turn structures. Grabowski et al. developed a virtual screening technique that represents a molecular scaffold by its side-chain attachment points (exit vectors) and properties of the side-chain substituents. They applied it for identification of β-turn mimetics and HMG-CoA inhibitors. Shuto et al. introduced principal component analysis (PCA) for the quantification of the three-dimensional (3D) structural diversity of mimetics in the chemical space and evaluated their stereoisomeric cyclopropane scaffolds in comparison with conformations of natural tetrapeptide motifs in an X-ray structure database. These studies demonstrate the utility of analytical elements for the structural analysis of peptide fragments in proteins.
Results of those earlier studies indicate the necessity of establishing a new method for visual and comprehensive evaluation of the conformations of the peptide fragments and peptidomimetics in the chemical space. Therefore, we attempted principal component analysis based on Cα–Cβ bond vectors combined with the alignment-free shape comparison method and ultrafast shape recognition (USR) encoding.35

Figure 3a presents calculation methods and the workflow for PCD-plot[0123]. Details of the procedures are described in Method 1 and Figure S1. Figure 4 presents the results of distribution of conformations of the extracted peptide fragments constituted by i, i+1, i+2, and i+3 amino acid residues from

Figure 2. Lists of structural peptidomimetics (a) single-facial mimetics and (b) double-facial and multifacial mimetics. Chemical structures, mimetic amino acids/motifs, target proteins and activities, the number of rotatable bonds involved in the scaffolds, and references are shown. Substituents highlighted in gray are "pseudo-Cα–Cβ bonds", which are designed to project side-chain Cα–Cβ bonds. The pseudo-Cα–Cβ bonds in these structural mimetics were assigned according to the description in the references. The i substituent of 5 and the i+4 substituent of 6 are not pseudo-Cα–Cβ bonds but pharmacophore mimetics (Figure 1 and its legend present the details). The rotational bonds are counted in the scaffolds inside the pseudo-Cα–Cβ bonds.
nonredundant sets of 118 proteins. The 28,105 extracted peptide fragment data set includes diverse structures such as the helix, sheet, turn, disordered, and random conformations and is appropriate for constructing the PCD plot, the chemical space of bond vectors. We designated this map as a peptide conformation distribution plot: PCD-plot. Numbers in square bracket notation represent the source of the side-chain number. Secondary structures were annotated for extracted peptide fragment motifs as the "Helix", "Sheet", "Turn", and "Others". Also, USR encoding was used to represent a shape formed by the set of $C_{\alpha}$-$C_{\beta}$ bond vectors extracted from a peptide fragment as a vector of 12 shape descriptors. Because USR is independent of the absolute coordinate system, molecular shapes can be compared directly without superposing the molecules. For this reason, the PCD plot is alignment free.

Figure 4 clarifies in a PCD plot that Helix (red), Turn (orange), and Sheet (green) form a cluster. The 9,372 Helix data gather on the left-side narrow area. The standard $\alpha$-helix (blue triangle) is located at the Helix center, indicating that the description of the $\alpha$-helix in the PCD plot is consistent with that in the Ramachandran plot because the standard helix is defined as a highly populated $\alpha$-helix area by $\phi$ and $\psi$ angles. Consequently, the PCD plot can analyze peptide secondary structures effectively in the chemical space and can distinguish...
secondary structures of peptide fragments using coordinates of Cα and Cβ atoms, instead of φ and ψ angles in a Ramachandran plot. Several examples of structures demonstrate that a partially distorted and unwound α-helix is positioned at the edge of or outside of the Helix region (Figure S2). Distribution of the Turn is overlapped with the Helix. In actuality, the conformations of Cα−Cβ bonds in the type I β-turn closely resemble those in the α-helix, except for the vector of the i+3 side chain (Figure S3).

Next, the multifacial mimetic scaffolds 10−12 were projected. They have i, i+1, i+2, and i+3 pseudo-Cα−Cβ bonds on PCD-plot[0123] (Figure 5). Calculation works are presented in Figure 3b and Method 2. Calculation examples are shown in Figure S4. This analysis, an alignment-free method, uses conformations of peptidomimetics projected to the PCD plot representing the chemical space of peptide fragments. Multiple conformations were assessed for each molecule using conformational search calculations. Accordingly, we can clarify the distribution of conformations of mimetic compounds and can also avoid arbitrary selection of a specified conformation suitable for the superposition. Black dots in Figure 5 denote each conformation of a scaffold. Conformations of 10 are located on the central region of the map. They are distributed widely because of high flexibility of the i side chain (Figure 5a). Conformations of 11 are spread out quite broadly, including the Helix region (Figure 5b) because 11 has six rotatable bonds in the scaffold and has high conformational flexibility. In this case, large entropic costs are concerned in the binding of the target protein. The conformation distribution of 12 is located on the limited area within the Helix region (Figure 5c), indicating 12 as a superior scaffold for multifacial α-helix mimetics.

The single-facial mimetics 1−4 were also projected on PCD-plot[047]. They have i, i+4, and i+7 pseudo-Cα−Cβ bonds (Figure 6a−d). To analyze a different set of pseudo-Cα−Cβ bonds, a separate PCD plot must be prepared for the corresponding Cα−Cβ bond sets. The calculation method is the same as that shown in Figure 4, where sets of the i, i+1, i+2, and i+3 sets in PCD-plot[0123]. The resulting maps and the meaning of the PCA-1 axis showed a similar tendency to that of PCD-plot[0123] (Table S3). The conformations of 1 and 2 are gathered in a small area. They are located near the left side of the Helix region, whereas those of 3 are located onto the Helix area. Conformations of 4 (with seven rotatable bonds) are spread broadly, but some parts are distributed within the Helix region. Scaffolds 5 and 6 cannot be projected on a PCD plot because some side chains are not pseudo-Cα−Cβ bonds but are pharmacophore mimetics (Figure 1b). Double-facial mimetics 7−9 having i, i+3, and i+4 pseudo-Cα−Cβ bonds were also projected on PCD-plot[034] (Figure 6e−g). Distributions of all mimetics are located on the limited area within or near the Helix region, indicating that these scaffolds are good α-helix mimetics.

In summary, the PCD plot is an alignment-free analysis based on the Cα−Cβ atom coordinates. It enables visual and comprehensive understanding of the conformations of peptide
fragments and classification of secondary structures likely by the Ramachandran plot but based on an alternative concept. Projection of a peptidomimetic molecule to the PCD plot visually clarifies the position of their conformation distribution.

Figure 6. (a−d) PCD-plot[047]: PCA analysis using the $i$, $i+4$, and $i+7$ side chains and projection of single-facial mimetic compounds 1−4. (e−g) PCD-plot[034]: PCA analysis using $i$, $i+3$, and $i+4$ side chains and projection of double-facial mimetic compounds 7−9. Notations are the same as those used in Figure 3.
and how widely their conformations spread on the chemical space of peptide fragments. It also reveals the three-dimensional structural similarity of mimetics to its target peptide and protein secondary structures. Peptidomimetic analysis using the PCD plot showed that mimetic scaffolds 7, 8, 9, and 12 are excellent \(\alpha\)-helix mimetics.

**PMA Map: Alignment Method Based on C\(_\alpha\)–C\(_\beta\) Bonds.** Projection of structural peptidomimetics on the PCD plot facilitates our understanding of their conformational features visually in the chemical space. However, from the prospective of drug design and scaffold selection, it is important to ascertain how precisely, in detail and quantitatively, the molecules can mimic a specified target motif. We therefore developed the peptidomimetic analysis map (PMA map), shown in Figure 7a. The calculation workflow is presented in Figure 3b. Detailed procedures are described in Method 3 and Figure S5. This

Figure 7. (a) Helix mimetics analyzer (HMA) map. The \(x\)-axis and \(y\)-axis, respectively, denote the average of position difference (APD, Å) and the average of vector difference (AVD). Error bars show the standard deviation. (b) Detailed helix mimetic analysis of 12 for each pseudo-C\(_\alpha\)–C\(_\beta\) bond. (c) Orientational distribution of the i pseudo-C\(_\alpha\)–C\(_\beta\) bond of 12. (d) Superposed views with \(\alpha\)-helix and mimetic compounds. Yellow and red denote each conformer and C\(_\alpha\)–C\(_\beta\) bond, respectively. Chemical structures, all conformers (left), and a representative conformer for clarification (right).
analysis is regarded as an alignment-based method because it is based on the structural superposition of a peptidomimetics and its target peptide motif. In this map, the x-axis is the average of position difference (APD), the root-mean-square deviation (RMSD, Å) between Cα atoms in a peptide motif and the corresponding atoms in a mimetic molecule. The y-axis is the average of vector difference (AVD), where the difference of orientation between a Cα−Cβ bond vector and its pseudo-Cα−Cβ bond vector is represented as “1−inner product of the two vectors”. When a standard α-helix is set to a reference motif, we designate it the helix mimetic analysis map (HMA map). Different from the PCD plot, the PMA map is independent of the choice of the set of side chains. In other words, we can compare peptidomimetics with different sets of pseudo-Cα−Cβ bonds, such as 1 and 12, on the same PMA map and can compare them directly.

Figure 7a clearly presents quantified analysis results of the structural mimetics shown in Figure 2. Scaffolds 9 and 12 have APD values of less than 0.5 Å in the map, meaning that the positions of the Cα-containing atoms in scaffolds 9 and 12 are closer to those of Cα atoms of the standard α-helix than those of others. The AVD values of scaffolds 9 and 12 are around 0.6, which means that the directions of their pseudo-Cα−Cβ bond vectors show better agreement with those of Cα−Cβ bond vectors of a standard α-helix than those of other mimetic scaffolds. The two scaffolds also showed good results on the PCD plot, as described above. Scaffolds 1, 2, 3, 4, 10, and 11 are more deviated from the standard α-helix in the HMA map, which are consistent with the distribution location in their PCD plot. Error bars of 4 and 11 in the map are large, which also accords with their wide distribution on the PCD plot and the crowded superposing view (Figure 7d). The HMA map shows conformity with the α-helix structure, with similarity between the Cα−Cβ bonds of the α-helix and their pseudo-Cα−Cβ bonds, which differs among mimetic scaffolds. The results are consistent with those of the PCD plot for most of the evaluated peptidomimetic scaffolds.

Detailed analyses of pseudo-Cα−Cβ bonds were performed with scaffold 12 (Figure 7b) and other scaffolds (Figure S6). In scaffold 12, the PD values of all pseudo-Cα−Cβ bonds are around 0.4 and good superposition with the four Cα atoms. The i + 2 VD value is 0.03 ± 0.03, which means that the i+2 pseudo-Cα−Cβ bond mimics the i+2 moiety of the standard α-helix almost completely. The i+1 VD value is 0.84 ± 0.29 and large, which is derived from the fact that the stereochemistry of the carbon atom at the 3-position of 12 is reverse to that of the natural amino acid. The i and i+3 VD values are, respectively, 0.66 ± 0.72 and 0.94 ± 0.38. These large deviations are attributable to the flip-flop of the carbamate group for the i VD value (Figure 7c) and existence of a free rotatable bond for the i + 3 VD value. These analyses of 12 revealed that the positions of the Cα′-corresponding atoms are close to those of the four Cα atoms of the α-helix and revealed that the mimetics has flexibility in the direction of its pseudo-Cα−Cβ bonds. Although this point can also be confirmed visually from superposed views (Figure 7d), quantitative analysis of each pseudo-Cα−Cβ bond explicitly shows characteristics such as the position, direction, flexibility, and sophistication of mimicking, which can guide the development or improvement of new mimetic scaffolds.

The PMA map, an alignment-based analysis, uses Cα−Cβ bonds and pseudo-Cα−Cβ bonds. It can evaluate the similarity of a mimetic compound with its target peptide motif in chemical space by quantified parameters. It incorporates consideration of the conformation and flexibility of the mimetic molecules. Moreover, it is useful to compare three-dimensional structures of different peptidomimetics having different sets of pseudo-Cα−Cβ bonds. The HMA map clarified that scaffolds 9 and 12 are α-helix mimetics superior to others.

Detailed Analysis of the PCD Plot. Here, we discuss the usefulness of the PCD plot based on the meanings of axes. Coefficients of the PCA-1 and PCA-2 axes are presented in Table S3. Moment-1 and moment-2 of fct and ftf of PCA-1 showed large coefficients in PCD-plot[0123], suggesting that PCA-1 roughly represents the molecular size. The coefficients of PCA-1 in PCD-plot[047] and PCD-plot[034] have similar characters. For example, scaffold 1 is projected on the left side of the Helix region (Figure 6a). The meaning of PCA-1 implies that the molecule is smaller than the α-helix. In actuality, a superposing view of the α-helix and 1 shows that 1 is shorter than the α-helix (Figure 7d). It is interesting that the alignment-based analysis and alignment-free analysis point out the same feature, but differently. The meanings of the PCA-2 axis differ among PCD plots; they are complicated. Specific examination of PCD-plot[0123] reveals that moment-3 occupies the major component, which represents the molecule asymmetry and skewness. Conformations on the lower area of the map show a symmetric shape, whereas those in the upper area have symmetry-breaking structures (Figure S7).

Regarding PCD-plot[0123], α-helix structures (shrink form regularly arranged by hydrogen bonds) are located on the left side of the map, whereas β-turn structures (extended form regularly arranged by hydrogen bonds) are located on the right side. In these areas, the distribution range on the PCA-2 axis is limited. In contrast, random structures (gray dots) range mainly on the middle region of the PCA-1 axis but spread widely along the PCA-2 axis. This result is consistent with the fact that various structures are included, such as symmetric and distorted conformations without hydrogen bond constraints. Therefore, the resulting chemical spaces in PCD-plot[0123] form a triangle.

The ratio of the explained variance of the PCA-1 is greater than 0.9. In the case of PCD-plot[0123], the 24 coordinate components extracted from four Cα−Cβ bonds of peptide fragments are ultimately reduced to two axes using principal component analysis. Alternatively, the relative positions of the four Cα−Cβ bonds can be expressed by eight φ and ψ angles. Moreover, the angles are not independent. As presented in the Ramachandran plot, the angles are limited in a specified set and are mutually related. Therefore, the chemical space of Cα−Cβ bond sets might be expressed intrinsically using a small set of independent variants. The PCD plot expresses the chemical space by Cα−Cβ bonds of peptide fragment motifs. The Ramachandran plot indicates peptide secondary structures by φ and ψ angles of each amino acid residue in peptides. Consequently, the PCD plot can be a useful and visible method for representing protein structures and peptidomimetics in a chemical space, similar to the Ramachandran plot, however, based on different analytical viewpoints.

Comparison of the PCD Plot and PMA Map. Table 1 presents the PCD plot and PMA map features. Both methods are based on analysis of Cα−Cβ bonds and pseudo-Cα−Cβ bonds, not φ and ψ angles, in which multiple conformations for mimetic molecules are considered. The PCD plot, an alignment-free method, can visualize the three-dimensional structure of peptide fragments comprehensively. Projection of a mimetic molecule clarifies its conformation distribution in the chemical space and
similarity with its target peptide secondary structures. We can infer that a molecule with a widespread conformational distribution would require entropic costs in binding to its target protein. It is noteworthy that the PCD plot can express even a deviated distribution of conformations because each dot represents a single conformation. We prepared PCD-plot[0123], PCD-plot[047], and PCD-plot[034], where a separate map must be used to compare mimetic molecules with different sets of pseudo-Cα-Cβ bonds.

The PMA map, in contrast, is an alignment-based method. Unlike the PCD plot, a single PMA map can compare various peptidomimetics with different sets of pseudo-Cα-Cβ bonds. The map can evaluate the similarity between an arbitrary target motif and its mimetic molecules using APD and AVD values. Unlike the PCD plot, this analysis provides detailed similarity indices between peptidomimetics. It gives quantification of the similarity of peptidomimetics based on features of the respective pseudo-Cα-Cβ bonds. The two methods clarified that 9 and our scaffold 12 are α-helix mimetics superior to others examined for this study. These results demonstrate that the two methods are complementarily useful for rational PPI drug discovery. Consequently, by applying these methods, we are now developing new peptidomimetic scaffolds, conducting PPI drug discovery, and constructing a compound library targeting PPIs, a subject that will be described in a future publication.

## METHODS

### Preparation of Nonredundant Protein Data Sets and the PCD Plot.

Workflows are shown in Figure 3a. Detailed procedures and examples for PCD-plot[0123] and PCD-plot[034] are presented in Figure S1.

Step A1. Preparation of nonredundant protein data sets. The selection method is based on a description presented in the literature. From conditions on X-ray diffraction to a resolution of 2.0 Å or higher, the number of distinct protein entities = 1, the sequence length >100, human proteins, and registration after the year 2000, we selected 7,178 protein structures from the Protein Data Bank (PDB) as on 27 August 2020. From the selected proteins, we extracted 200 proteins randomly and removed redundant proteins by manual inspection. The final sets of selected structures were 118 nonredundant proteins containing 28,764 amino acids (List S1). The DSSP secondary structure types were assigned for each amino acid. The number of the respective DSSP secondary structure types is listed in Table S1.

Step A2. Extraction of peptide fragments. To make PCD-plot[0123], the “i, i+1, i+2, and i+3” peptide fragment motifs were extracted from the head to the tail with one amino acid shift from the selected 118 proteins. The number of extracted motifs was 28,105. Examples of fragments 1–5 from PDB ID:1g60 are shown in Figure S1. Secondary structures “Helix”, “Sheet”, “Turn”, and “Others” were annotated for each motif. When the DSSP types of four amino acids have three or four continuous G/H/I, T, and E, the motifs are defined, respectively, as Helix, Turn, and Sheet. The remaining types were assigned to Others. For PCD-plot[047] and PCD-plot[034], the “i, i+4, and i+7” and the “i, i+3, and i+4” motifs were extracted, respectively. Secondary structures were annotated similarly. In these two motifs, Turn is not defined because the β-turn structure is constructed by four continuous amino acids. Table S2 lists the number of annotated secondary structures for the extracted motifs (peptide fragments).

Step A3. Extraction of the Cα and Cβ atom coordinates from each fragment. The XYZ coordinates of “i, i+1, i+2, and i+3” Cα and Cβ atoms are extracted for PCD-plot[0123]; the XYZ coordinates of “i, i+3, and i+4” Cα and Cβ atoms are extracted for PCD-plot[034]. Examples are shown in Figure S1.

Step A4. Conversion to ultrafast shape recognition (USR) descriptors. The extracted coordinate sets were converted to a vector of 12 shape descriptors using USR encoding, which is calculated from the distribution of the interatomic distance derived from a set of four reference points. As an alignment-free shape comparison method, USR has been used in virtual screening against arylamine N-acetyltransferases. The 12 descriptors used for calculations consist of four molecular locations and three moments. The four molecular locations are the molecular centroid (cld), the closest atom to cld (cst), the farthest atom from cld (fct), and the farthest atom from fct (fft). As the three moments, the first moment is the average atomic distance to the molecular centroid (an estimate of the molecular size). The second is the square root of the variance of these
atomic distances about the first moment. The third moment is the skewness of these atomic distances about the first moment (i.e., a measure of the asymmetry of the distribution). We used the skewness instead of the cubic root of the skewness used in the original report from the literature because we balanced the descriptors to make a better classification of secondary structures of peptide fragment motifs. The balances of the roots are also discussed in that original report.

Steps A5 and A6. Dimension reduction was conducted using principal component analysis (PCA) to generate a PCD plot. The first and second components (PCA-1 and PCA-2) were assigned to x- and y-axes for 2D graph representation. The PCA calculations were performed using Scikit-learn (https://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html).

Conformation Generation of Peptidomimetics and Its Projection to the PCD Plot. Figure 3b portrays a workflow for the conformational search of peptidomimetics and its projection to the PCD plot. Details of the procedures and examples of scaffold 12 are displayed in Figure S4 for clarification.

Step B1. Conformational search of peptidomimetics. Conformations of mimetic molecules were generated using the “Conformer Distribution” function of computational chemistry software Spartan’18 ver. 1.4.4 (Wavefunction Inc.). The calculations were conducted using the MMFF force field and “SEARCHMETHOD = MONTECARLO, FINDBOATS” options, which can explore a wide conformation space including the flip-flop of scaffold rings. All side chains in mimetic molecules are modeled to a methyl group (corresponding to the alanine side chain) to avoid unsuccessful side-chain conformation searching. Considering the coarse-grained calculation level, all generated conformers within 10 kcal/mol from the most stable conformation were adopted for analyses.

The free rotatable bonds in the peptidomimetic molecules presented in Figure 2 are few (2–7). Therefore, most of the possible conformational space can be covered by the conformational search using Spartan software. Hehre et al. report the validity of conformational ensemble using Spartan based on agreement of NMR chemical shifts calculated from the generated conformers with the experimental data. Importantly, PCD plot and PMA map results on peptidomimetic molecules depend on the obtained conformational ensemble. Therefore, in the analysis of larger and more flexible peptidomimetic molecules, another conformational search method such as multicanonical MD might be appropriate.

Step B2. Generated conformations were exported to sdf files. The XYZ coordinate sets of pseudo-Cα–Cβ bonds were extracted from each conformer. For scaffold 12, the coordinates of “i, i+1, i+2, and i+3” pseudo-Cα and Cβ atoms are extracted (Figure S4). This step corresponds to step A3.

Step B3. The 24 extracted coordinate data sets were converted to a vector of 12 shape descriptors, as described in step A4.

Step B4. The PCA-1 and PCA-2 values were calculated using the PCA coefficients (Table S3) found in step A5. They are projected to the PCD plot.

Superposition with Standard α-Helix and PMA Map Generation. Figure 3b presents a workflow for the procedures (steps B5 and B6). Details of the procedures and examples of scaffold 12 are shown in Figure S5.

Steps B1 and B2 are the same as those in Method 2.

Step B5. Superposition to a target motif. Cα atoms (pseudo-Cα atoms) in pseudo-Cα–Cβ bonds of a mimetic molecule were superimposed onto the corresponding Cα atoms of a target motif using the rdAlignment module implemented in RDKit: Open-source cheminformatics software (https://www.rdkit.org/). In mimetic scaffold 12, four pseudo-Cα atoms were superposed onto the four Cα atoms in a standard α-helix. The standard α-helix was generated using the poly-Ala sequence with ϕ = −63.8° and ψ = −41.1°, which is known as the highly populated area of the actual α-helix in the PDB.

Step B6. Calculation of position and vector differences and projection to the PMA map. The position difference (PD, Å) is defined as the Euclidean distance between the pseudo-Cα atom of a molecule and the corresponding Cα atom of a target motif. The APD is defined as RMSD between the pseudo-Cα atoms of a mimetic molecule and the corresponding Cα atoms of a target motif. In addition, after the superposition, the vector difference (VD) is calculated as

$$VD = 1 - \langle v_{α}^p, v_{α}^M \rangle$$

where $v_{α}^p$ is a unit vector from the Cα coordinate to the Cβ coordinate of a target motif, $v_{α}^M$ is a unit vector from the Cα coordinate to the Cβ coordinate of the pseudo-Cα–Cβ bond in a mimetic molecule, and the inner product is denoted by $\langle \cdot, \cdot \rangle$. In this formulation, each of the unit vectors was translated so that its starting point (Cα coordinate) was at the origin. The average value of VD for every conformation is defined as AVD. The VD value is normalized. It ranges from 0 to 2. The APD/AVD values and their standard deviations used for the PMA map were calculated from every conformer and every pseudo-Cα–Cβ bond in a mimetic molecule.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.1c03967.

PDB ID of 118 nonredundant proteins used for analyses; detailed procedures and examples for the PCD plot; assignment of DSSP secondary structure types; secondary structure annotation for extracted peptide fragment motifs; coefficients of PCA axes; examples of peptide fragment structures in the PCD plot; superposed views of the α-helix and β-turn; detailed procedures and examples for mapping conformations to the PCD plot; detailed procedures and examples for superposition and PMA-map generation; detailed helix mimetic analysis of mimetic scaffolds 1–11; and examples of peptide fragment conformers (PDF)

AUTHOR INFORMATION

Corresponding Author

Hajime Takashima — Research and Development Department, PRISM BioLab Co., Ltd., Fujisawa, Kanagawa 251-0012, Japan; orcid.org/0000-0002-7436-040X; Phone: +81-466-25-2555; Email: takashima@prismbiolab.com

Authors

Atsushi Yoshimori — Chemoinformatics & AI Research Group, Institute for Theoretical Medicine, Inc., Fujisawa, Kanagawa 251-0012, Japan

Eiji Honda — Research and Development Department, PRISM BioLab Co., Ltd., Fujisawa, Kanagawa 251-0012, Japan

Tomonori Taguri — Research and Development Department, PRISM BioLab Co., Ltd., Fujisawa, Kanagawa 251-0012, Japan
Junchan Ozawa — Research and Development Department, PRISM BioLab Co., Ltd., Fujisawa, Kanagawa 251-0012, Japan
Masaji Kasai — Research and Development Department, PRISM BioLab Co., Ltd., Fujisawa, Kanagawa 251-0012, Japan
Satoshi Shuto — Faculty of Pharmaceutical Science, Hokkaido University, Sapporo, Hokkaido 060-0812, Japan; orcid.org/0000-0001-7850-8064
Dai Takehara — Research and Development Department, PRISM BioLab Co., Ltd., Fujisawa, Kanagawa 251-0012, Japan

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.1c03967

Author Contributions
H.T. and A.Y. conceived and designed the research. H.T., A.Y., E.H., T.T., J.O., and D.T. performed the research. H.T. and A.Y. wrote the article. E.H. and M.K. revised the article. S.S. proposed the name of the pseudo-Cα-Cβ bond and revised the article. D.T. supervised the research. All authors have given approval to the final version of the manuscript.

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Dr. Hiroyuki Kouji for contributions on the prototype of the PMA map.

■ ABBREVIATIONS

APD, average of position difference; AVD, average of vector difference; HMA, helix mimic; NRSF/REST, neuron-restrictive silencer factor/RE1-silencing transcription factor; PCA, principal component analysis; PCD, peptide conformational distribution; PD, position difference; PDB, Protein Data Bank; PMA, peptidomimetic analysis; PPIs, protein–protein interactions; RMSD, root-mean-square deviation; USR, ultrafast shape recognition; VD, vector difference

■ REFERENCES

(1) Conte, L. L.; Chothia, C.; Janin, J. The Atomic Structure of Protein-Protein Recognition Sites. J. Mol. Biol. 1999, 285, 2177–2198.
(2) Guharoy, M.; Chakrabarti, P. Secondary Structure Based Analysis and Classification of Biological Interfaces: Identification of Binding Motifs in Protein-Protein Interactions. Bioinformatics 2007, 23, 1909–1918.
(3) Sawyer, N.; Watkins, A. M.; Arora, P. S. Protein Domain Mimics as Modulators of Protein–Protein Interactions. Acc. Chem. Res. 2017, 50, 1313–1322.
(4) Milroy, L. G.; Grossmann, T. N.; Sven, H.; Luc, B.; Christian, O. Modulators of Protein-Protein Interactions. Chem. Rev. 2014, 114, 4695–4748.
(5) Bullock, B. N.; Jochem, A. L.; Arora, P. S. Assessing Helical Protein Interfaces for Inhibitor Design. J. Am. Chem. Soc. 2011, 133, 14220–14223.
(6) Pelay-Gimeno, M.; Glas, A.; Koch, O.; Grossmann, T. N. Structure-Based Design of Inhibitors of Protein-Protein Interactions: Mimicking Peptide Binding Epitopes. Angew. Chem., Int. Ed. 2015, 54, 8896–8927.
(7) Watkins, A. M.; Bonneau, R.; Arora, P. S. Side-Chain Conformational Preferences Govern Protein-Protein Interactions. J. Am. Chem. Soc. 2016, 138, 10386–10389.
(8) Davis, J. M.; Tsou, L. K.; Hamilton, A. D. Synthetic Non-Peptide Mimetics of α-Helices. Chem. Soc. Rev. 2007, 36, 326–334.
(9) Jayatunga, M. K. P.; Thompson, S.; Hamilton, A. D. α-Helix Mimetics: Outwards and Upwards. Bioorg. Med. Chem. Lett. 2014, 24, 717–724.
(10) Wilson, A. J. Helix Mimetics: Recent Developments. Prog. Biophys. Mol. Biol. 2015, 119, 33–40.
(11) Yin, H.; Lee, G. I.; Hyung, S. P.; Payne, G. A.; Rodriguez, J. M.; Sebti, S. M.; Hamilton, A. D. Terphenyl-Based Helical Mimetics That Disrupt the P53/MDM2 Interaction. Angew. Chem., Int. Ed. 2005, 44, 2704–2707.
(12) Kutzki, O.; Park, H. S.; Ernst, J. T.; Orner, B. P.; Yin, H.; Hamilton, A. D. Development of a Potent Bcl-2 Antagonist Based on α-Helix Mimicry. J. Am. Chem. Soc. 2002, 124, 11838–11839.
(13) Yin, H.; Lee, G. I.; Sedey, K. A.; Rodriguez, J. M.; Wang, H. G.; Sebti, S. M.; Hamilton, A. D. Terephthalamidine Derivatives as Mimetics of Helical Peptides: Disruption of the Bcl-2/Bak Interaction. J. Am. Chem. Soc. 2005, 127, 5463–5468.
(14) Shaginian, A.; Whitby, L. R.; Hong, S.; Hwang, I.; Farooqui, B.; Searcey, M.; Chen, J.; Vogt, P. K.; Boger, D. L. Design, Synthesis, and Evaluation of an α-Helix Mimetic Library Targeting Protein-Protein Interactions. J. Am. Chem. Soc. 2009, 131, 5564–5572.
(15) Burslem, G. M.; Kyle, H. F.; Breeze, A. L.; Edwards, T. A.; Nelson, A.; Warriner, S. L.; Wilson, A. J. Small-Molecule Peptideomimetic Inhibitors of the HIF-1α-P300 Protein-Protein Interaction. ChemBioChem 2014, 15, 1083–1087.
(16) Moon, H.; Lee, W. S.; Oh, M.; Lee, H.; Lee, J. H.; Im, W.; Lim, H. S. Design, Solid-Phase Synthesis, and Evaluation of a Phenyl-Piperazine-Triazine Scaffold as α-Helix Mimetics. ACS Comb. Sci. 2014, 16, 695–701.
(17) Tošovská, P.; Arora, P. S. Oligooxopiperazines as Nonpeptidic α-Helix Mimetics. Org. Lett. 2010, 12, 1588–1591.
(18) Lao, B. B.; Drew, K.; Guarracino, D. A.; Brewer, T. F.; Heindel, D. W.; Bonneau, R.; Arora, P. S. Rational Design of Topographical Helix Mimics as Potent Inhibitors of Protein-Protein Interactions. J. Am. Chem. Soc. 2014, 136, 7787–7788.
(19) Rodriguez, A. L.; Tamrazi, A.; Collins, M. L.; Katzenellenbogen, J. A. Design, Synthesis, and in Vitro Biological Evaluation of Small Molecule Inhibitors of Estrogen Receptor α Coactivator Binding. J. Med. Chem. 2004, 47, 600–611.
(20) Becerril, J.; Hamilton, A. D. Helix Mimetics as Inhibitors of the Interaction of the Estrogen Receptor with Coactivator Peptides. Angew. Chem., Int. Ed. 2007, 46, 4471–4473.
(21) Jacoby, E. Biphenyls as Potential Mimetics of Protein α-Helix. Bioorg. Med. Chem. Lett. 2002, 12, 891–893.
(22) Golebiowski, A.; Klopfenstein, S. R.; Chen, J. J.; Shao, X. Solid Supported High-Throughput Organic Synthesis of Peptide β-Turn Mimetics via Petasis Reaction/Diketopiperazine Formation. Tetrahedron Lett. 2000, 41, 4841–4844.
(23) Mizuno, A.; Kameda, T.; Kuwahara, T.; Endoh, H.; Ito, Y.; Yamada, S.; Hasegawa, K.; Yamano, A.; Watanabe, M.; Arisawa, M.; Shuto, S. Cyclopropane-Based Peptidomimetics Mimicking Wide- ranging Secondary Structures of Peptides: Conformational Analysis and Their Use in Rational Ligand Optimization. Chem. – Eur. J. 2017, 23, 3159–3168.
(24) Mizuno, A.; Matsui, K.; Shuto, S. From Peptides to Peptidomimetics: A Strategy Based on the Structural Features of Cyclopropane. Chem. – Eur. J. 2017, 23, 14394–14409.
(25) Ueda, H.; Kurita, J.; Neyaama, H.; Hirao, Y.; Kouji, H.; Mishina, T.; Kasai, M.; Nakano, H.; Yoshimori, A.; Nishimura, Y. A Mimetic of the MSin3-Binding Helix of NRSF/REST Ameliorates Abnormal Pain Behavior in Chronic Pain Models. Bioorg. Med. Chem. Lett. 2017, 27, 4705–4709.
(26) Schoenherr, C. J.; Anderson, D. J. The Neuron-Restrictive Silencer Factor (NRSF): A Coordinate Repressor of Multiple Neuron-Specific Genes. Science 1995, 267, 1360–1363.
(27) Chong, J. A.; Tapia-Ramirez, J.; Kim, S.; Toledo-Aral, J. J.; Zheng, Y.; Boutros, M. C.; Altschuller, Y. M.; Frohman, M. A.; Kramer, S. D.; Mandel, G. REST: A Mammalian Silencer Protein That Restricts Sodium Channel Gene Expression to Neurons. Cell 1995, 80, 949–957.
(28) Nomura, M.; Uda-Tochio, H.; Murai, K.; Mori, N.; Nishimura, Y. The Neural Repressor NRSF/REST Binds the PAH1 Domain of the Sin3 Corepressor by Using Its Distinct Short Hydrophobic Helix. *J. Mol. Biol.* 2005, 354, 903–915.
(29) Kawase, H.; Ago, Y.; Naito, M.; Higuchi, M.; Hara, Y.; Hasebe, S.; Tsukada, S.; Kasai, A.; Nakazawa, T.; Mishina, T.; Kouji, H.; Takuma, K.; Hashimoto, H. MS-11, a Mimetic of the MSin3-Binding Helix in NRSF, Ameliorates Social Interaction Deficits in a Prenatal Valproic Acid-Induced Autism Mouse Model. *Pharmacol. Biochem. Behav.* 2019, 176, 1–5.
(30) Higo, J.; Takashima, H.; Fukunishi, Y.; Yoshimori, A. Generalized-Ensemble Method Study: A Helix-Mimetic Compound Inhibits Protein-Protein Interaction by Long-Range and Short-Range Intermolecular Interactions. *J. Comput. Chem.* 2021, 42, 956–969.
(31) Ramachandran, G. N.; Sasisekharan, V. Conformation of Polypeptides and Proteins. *Adv. Protein Chem.* 1968, 23, 283–437.
(32) Hovmöller, S.; Zhou, T.; Ohlson, T. Conformations of Amino Acids in Proteins. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2002, 58, 768–776.
(33) Garland, S. L.; Dean, P. M. Design Criteria for Molecular Mimics of Fragments of the $\beta$-Turn. 2. $\alpha$-$\beta$ Bond Vector Analysis. *J. Comput.-Aided. Mol. Des.* 1999, 13, 485–498.
(34) Grabowski, K.; Proschak, E.; Baringhaus, K.; Rau, O.; Schubert-Xsilavecs, M.; Schneider, G. Bioisosteric Replacement of Molecular Scaffolds: From Natural Products to Synthetic Compounds. *Nat. Prod. Commun.* 2008, 3, 1355–1360.
(35) Ballester, P. J.; Finn, P. W.; Richards, W. G. Ultrafast Shape Recognition: Evaluating a New Ligand-Based Virtual Screening Technology. *J. Mol. Graphics Model.* 2009, 27, 836–845.
(36) Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* 1983, 22, 2577–2637.
(37) Touw, W. G.; Baakman, C.; Black, J.; Te Beek, T. A. H.; Krieger, E.; Joosten, R. P.; Vriend, G. A Series of PDB-Related Databases for Everyday Needs. *Nucleic Acids Res.* 2015, 43, D364–D368.
(38) Ballester, P. J.; Westwood, I.; Laurieri, N.; Sim, E.; Richards, W. G. Prospective Virtual Screening with Ultrafast Shape Recognition: The Identification of Novel Inhibitors of Arylamine N-Acetyltransferases. *J. R. Soc. Interface* 2010, 7, 335–342.
(39) Hehre, W.; Klunzinger, P.; Deppmeier, B.; Driessen, A.; Uchida, N.; Hashimoto, M.; Fukushi, E.; Takata, Y. Efficient Protocol for Accurately Calculating $^{13}$C Chemical Shifts of Conformationally Flexible Natural Products: Scope, Assessment, and Limitations. *J. Nat. Prod.* 2019, 82, 2299–2306.