ORIGINAL ARTICLE

Deciphering the conformational landscape of few selected aromatic noncoded amino acids (NCAAs) for applications in rational design of peptide therapeutics

Lalita Mohan Behera[1] · Manaswini Ghosh[1] · Soumendra Rana[1]

Abstract

Amino acids are the essential building blocks of both synthetic and natural peptides, which are crucial for biological functions and also important as biological probes for mapping the complex protein–protein interactions (PPIs) in both prokaryotic and eukaryotic systems. Mapping the PPIs through the chemical biology approach provides pharmacologically relevant peptides, which can have agonistic or antagonistic effects on the targeted biological systems. It is evidenced that ≥ 60 peptide-based drugs have been approved by the US-FDA so far, and the number will improve further in the foreseeable future, as ≥ 140 peptides are currently in clinical trials. However, natural peptides often require fine-tuning of their pharmacological properties by strategically replacing the α_L-amino acids of the peptides with non-coded amino acids (NCAA), for which codons are absent in the genetic code for biosynthesis of proteins, prior to their applications as therapeutics. Considering the diverse repertoire of the NCAAs, the conformational space of many NCAAs is yet to be explored systematically in the context of the rational design of therapeutic peptides. The current study deciphers the conformational landscape of a few such Cα-substituted aromatic NCAAs (Ing: 2-indanyl-L-Glycine; Bpa: 4-benzoyl-L-phenylalanine; Aic: 2-aminoindane-2-carboxylic acid) both in the context of tripeptides and model synthetic peptide sequences, using alanine (Ala) and proline (Pro) as the reference. The combined data obtained from the computational and biophysical studies indicate the general success of this approach, which can be exploited further to rationally design optimized peptide sequences of unusual architecture with potent antimicrobial, antiviral, gluco-regulatory, immunomodulatory, and anti-inflammatory activities.

Keywords Non-coded amino acids · Peptide design · Peptide therapeutics · Circular dichroism · Fluorescence

Introduction

Amino acids are the primary building blocks of all proteins involved in various biological processes in living organisms. Twenty-three α-amino acids are commonly found in proteins, from which twenty α-amino acids are called proteinogenic, for which codons are available in the genetic code for biosynthesis of proteins. The remaining three are post-translationally modified α-amino acids, sometimes found only in matured proteins. Interestingly, ≥ 140 amino acids

are found in natural proteins due to the post-translational modifications of the naturally occurring twenty amino acids (Ambrogelly et al. 2007). For unclear reasons, nature has always chosen α_L-amino acids, glycine being the achiral exception, in the biosynthesis of proteins. However, many α_L and non-α_L-amino acids are also found in the living organisms for which codons are absent in the genetic code for protein synthesis. Such amino acids are collectively classified as non-proteinogenic/unnatural/non-coded amino acids (NCAAs) (Bell et al. 2008; Blaskovich 2016). Examples of a few such NCAAs have been listed below in Fig. 1.

Notably, ≥ 1000 NCAAs are naturally found in various plants and microorganisms (Bell 2003; Lu and Freeland 2006; Pollegioni and Servi 2012). The existence of NCAAs in living organisms has immensely inspired chemical biologists over the years for peptidomimetic manipulation, which has expanded the chemical and conformational landscapes of the native α-amino acids for potential usage in biological

✉ Soumendra Rana
soumendra@iitbbs.ac.in

1 Chemical Biology Laboratory, School of Basic Sciences, Indian Institute of Technology Bhubaneswar, Bhubaneswar, Odisha 752050, India
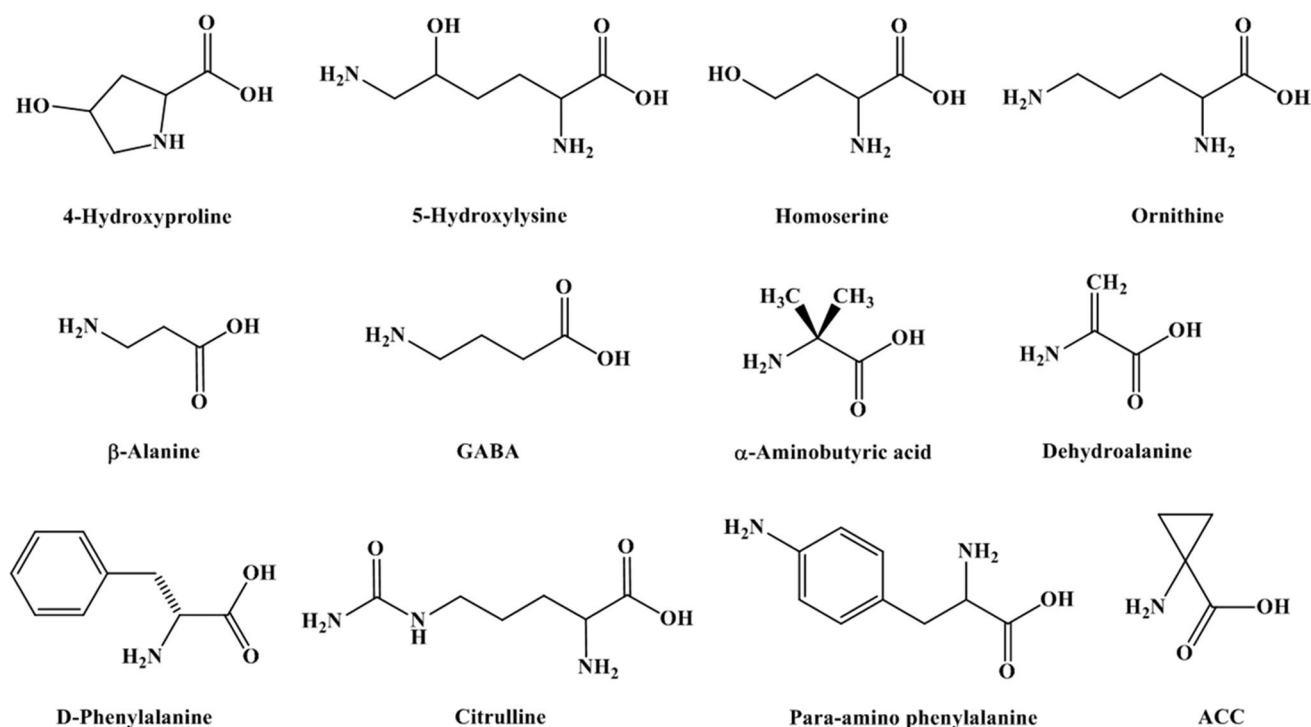
Fig. 1 Non-coded amino acids (NCAAs) commonly found in living organisms

applications. Interestingly, several NCAAs have been put to the test for enhancing the target-specific biological activity compared to their natural counterparts (Blaskovich 2016). Further, NCAAs can have numerous structural diversity and functional versatility and, thus, are often used as building blocks and molecular scaffolds in constructing combinatorial libraries (Goodman et al. 2007; Johnson et al. 2010). In addition, NCAAs are often incorporated into the peptides to improve the specific backbone conformational propensity and enhance the overall structural stability of the peptides (Balaram 1992; Rana et al. 2005; Vasudev et al. 2011). Indeed, the usage of NCAAs appears to be advantageous in various applications ranging from foldamer design and advanced material to peptide-based drugs (Clerici et al. 2016; Gellman 1998; Li and Yang 2006). NCAAs are widely used in peptidomimetics, and pharmacologically active analogs of the native peptides, as they confer better bioavailability and thermostability. No wonder that many of these NCAAs are also critical components in pharmaceuticals and developmental drugs (Blaskovich 2016) such as antitumor, antimicrobial, osteoporosis, diabetes, and hypertension (Avan et al. 2014; Vlieghe et al. 2010).

In the world of pharmacopeia, peptides as therapeutics bridge the gap between the traditional small molecules and the biologics, as it has a distinct biochemical and therapeutic profile. With ≥ 60 FDA-approved drugs in the market, peptides as therapeutics had a glorious past, and current trends

of ≥ 140 peptides in clinical trials suggest that peptides also have a fabulous future (Fosgerau and Hoffmann 2015). In general, approval of peptides as targeted therapeutics has a better success rate compared to the traditional small molecules, as peptides are generally known to be physiologically active and efficacious signaling molecules that can modulate cellular signaling by selectively binding to a plethora of transmembrane proteins like G-protein coupled receptors (Davenport et al. 2020) (GPCR) and ion channels, including some enzymes (Fosgerau and Hoffmann 2015; Lau and Dunn 2018). In addition, by the nature of their chemical constituent, peptides are chemically diverse and can be manufactured with ease due to established chemistry. Peptides are accepted to be safer as therapeutics as they have a relatively low immunogenicity profile compared to small molecules (Purcell et al. 2007). More importantly, peptides can selectively target larger surface areas typically involved in protein–protein interactions due to their intermediate size. From the therapeutic standpoint, natural peptides provide high levels of biodiversity, biological activity, specificity, low immunogenicity, and toxicity in therapeutic usages. However, peptides made of natural amino acids tend to have low metabolic stability due to the susceptibility toward proteolysis, leading to a short half-life, poor epithelial absorption, and rapid elimination from the body (Lien and Lowman 2003; Antosova et al. 2009). Altogether, it all adds up to the current bottlenecks of the peptide drug market, which is

mainly limited to injectables. Though many of the shortcomings associated with natural peptides are being addressed by the designer synthetic peptides, the oral bioavailability of the peptides remains a challenge due to the complex chemical-biological landscape of the gastrointestinal tract. In principle, the utilization of NCAAs in the rational design of therapeutic peptides (Fosgerau and Hoffmann 2015) can substantially overcome some of the limitations encountered by the natural peptides, enhancing their potential for therapeutic usage (Stevenazzi et al. 2014; Cavaco et al. 2018; Stone and Deber 2017). In addition, with the help of evolving formulations and drug delivery platforms, the scope of the peptide drug administration can expand beyond the traditional injectables to oral, intranasal, and transdermal routes.

The issue of bioavailability is currently being addressed by several strategies, out of which rational introduction of NCAAs to the natural peptide backbone appears to be quite promising to improve the oral bioavailability of the peptides (Ding et al. 2020). The presence of NCAAs can influence the membrane permeability and improve the binding affinity of the peptides toward the target proteins by inducing desired conformational changes. In addition, strategic positioning of the NCAAs in the peptide sequences through rational design can also introduce resistance to degradation in the gastrointestinal tract and improve plasma half-life, absorption, and immunogenicity profile of the peptide therapeutics (Di 2015; Knudsen et al. 2000). For example, stereochemical mutation (Durani 2008) of L to D-amino acids has been shown to substantially improve the pharmacokinetic property of octreotide compared to somatostatin, a drug prescribed for treating gastrointestinal tumors (Katz and Erstad 1989; Mahalakshmi and Balaram 2006). Further, simple chemical modification of the terminal amino acids (Rader et al. 2018) like *N*-acetylation and *N*-amidation, including *N*-methylation of amides, has been shown to improve resistance to proteolysis and delayed renal clearance in the case of cyclosporin, a drug prescribed to tame the immune system during organ transplant. Similarly, incorporating D-amino acids and NCAA like 2-naphthyl alanine through rational design has produced LY2510924 peptide, a potent CXCR4 receptor antagonist for treating tumors (Peng et al. 2015). In addition, the introduction of hydrocinnamic acid to PMX53, a potent cyclic peptide antagonist of the C5a receptor (C5aR1) (Liu et al. 2018; Rana et al. 2016), has been shown to improve its oral bioavailability substantially in animal studies (Kumar et al. 2020). Further, the presence of NCAAs like Aib has been shown to significantly improve the therapeutic index of glucagon-like peptide-1 (GLP-1) mimetics in the treatment of type-II diabetes (T2DM), including several other peptide-based drugs (Brunel et al. 2019; Sebokova et al. 2010). Similarly, the activity of the compstatin, a potent cyclic peptide inhibitor

of the complement protein C3, has been shown to improve remarkably by introducing 1-methyl tryptophan at position 4 as a replacement for the valine (Katragadda et al. 2006). In addition, the usage of α, β-dehydrophenylalanine in the rational design of peptides has also been shown to improve the antimicrobial activity of the short peptides (Mathur et al. 2007; Sharma et al. 2012). Further, incorporation of the NCAAs like L-ornithine and L-kynurenine into the antibiotics like daptomycin, often prescribed to treat life-threatening bacterial infections, indicates the untapped potential of the NCAAs (Miller et al. 2016; Gray and Wenzel 2020).

Nevertheless, several synthetically modified NCAAs are still there in the literature (Blaskovich 2016), whose basic functionality as a building block of designer peptides or foldamers has not been systematically explored to evaluate their scope in the future biotechnological applications. Interestingly, several NCAAs in free form and embedded in peptides are often found as the central component of many marketed drugs (Gupta and Chauhan 2011). It is also noteworthy that the NCAAs with aromatic structures are often incorporated into proteins through biosynthetic pathways. In general, NCAAs with mono and disubstituted bulky aromatic side chains at the Cα-position tend to be conformationally constrained, which is considered disadvantageous for racemization and sterically challenging for proteolysis. Therefore, the inclusion of such NCAAs to the peptide backbone improves the bioavailability of the natural peptides (Ramesh et al. 2011).

In this context, the current study has attempted to systematically explore the backbone conformational propensities of a few aromatic NCAAs (Fig. 2), such as 2-indanyl-L-Glycine (Ing), 4-benzoyl-L-phenyl alanine (Bpa), and 2-amino indane-2-carboxylic acid (Aic) for potential application in de novo design of bioactive peptides (Saghyan and Langer 2016), by recruiting the canonical amino acids like alanine (Ala) and proline (Pro) as the reference. Among the selected NCAAs, Bpa is a photo labeling aromatic amino acid, which has been utilized to study PPIs (Chin et al. 2002; Grunbeck et al. 2011; Joiner et al. 2019; Kauer et al. 1986). Further, both Ing and Aic are aromatic amino acids. Between the two, Aic appears to be a highly conformationally constrained sterically hindered amino acid, which is found in several drug-like molecules (Schiller et al. 1991; Kowalczyk et al. 2007). The computational and experimental data presented in the current study for the selected NCAAs in the context of model synthetic peptides indicate that the said amino acids (Ing, Bpa, and Aic) can be systematically utilized further in rational engineering of conformationally ordered biologically active small designer peptides with higher thermostability and bioavailability in a sequence-dependent manner.
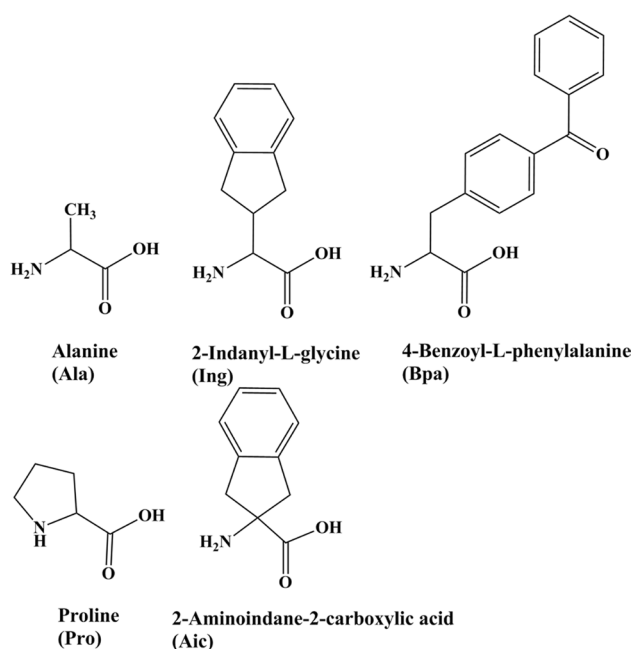
Fig. 2 Structural comparison of the naturally coded amino acids with the non-coded amino acids explored in the current study

Results

Conformational propensities of the selected NCAAs

Conformational propensities of the individual amino acids are usually dependent on the peptide sequence and the overall molecular environment imparted by the solvent conditions. Interestingly, the literature favors the quantum mechanical (QM) method (Revilla-Lopez et al. 2010; Bisetty et al. 2005; Tran et al. 2006) over the classical approach for reliably predicting the conformational propensities of both coded and NCAAs. Though the QM-based methods are relatively more accurate, the accuracy level is nearly similar to the molecular mechanics (MM) based methods for larger systems. In addition, the QM approach is computationally resource-intensive, even for small peptide systems in the presence of explicit water molecules. On the other hand, the computational cost required for classical MM-based molecular dynamics studies in the presence of explicit water molecules is comparatively minimal and has been successful in providing quality data that are relatable to the experimental observations (Croitoru et al. 2021; Beck et al. 2008). Thus, in the current study classical approach has been put to use for garnering the information about the accessible conformational space available to the selected NCAAs (Fig. 2), using the canonical coded amino acids as a reference in a model Ala-X-Ala (X = amino acid of interest) tripeptide system. Initially, Ala-Gly-Ala tripeptide, both acylated and amidated respectively, at the N- and the C-terminus was modeled in an extended conformation

$(\phi = -120°, \psi = 120°)$ and was subsequently subjected to molecular dynamics (MD) studies over 10 ns at 300 K in explicit SPC water for probing whether the 10 ns time would be enough for the model tripeptides to reasonably access the available conformational space in the four quadrants of the Ramachandran plot (Fig. S1).

Further, the Ac-Ala-**Ala**-Ala-NH$_2$, Ac-Ala-**Aib**-Ala-NH$_2$, and Ac-Ala-**Pro**-Ala-NH$_2$ tripeptides were also subjected to 10 ns of MD, and the backbone dihedral angles of the highlighted amino acids were further plotted at an interval of 10 ps. The conformational propensity data presented for the coded amino acids in Fig. 3 (Table 1) strongly correlate with the observations made in the structural studies of the proteins and peptides. Thus, the Ac-Ala-**Ing**-Ala-NH$_2$, Ac-Ala-**Bpa**-Ala-NH$_2$, and Ac-Ala-**Aic**-Ala-NH$_2$ tripeptides, containing the non-coded amino acids like 2-indanyl-L-Glycine, 4-benzoyl-L-phenyl alanine, and 2-amino indane-2-carboxylic acid were also subjected to 10 ns of MD at 300 K in the presence of explicit water.

In addition, the data presented for these amino acids in Fig. 3 and Table 1 indicate that both Ing and Bpa have conformational propensities broadly similar to Ala, whereas the Aic has a conformational propensity similar to Aib. It is noteworthy that while Aib has a hydrophobic side chain, Aic has a side chain that is both hydrophobic and aromatic in nature. Further, it is hypothesized that free Ing can exhibit both "endo" and "exo" conformations due to the possible pucker in the adjacent five-membered ring. However, quantum mechanical studies suggest that the "endo" conformer has lower energy than the "exo" conformer (Renfrew et al. 2012). Thus, Ing was modeled in the "endo" conformer. Interestingly, Ing in "endo" or "exo" conformer did not sample the right-handed helical conformations in the model tripeptide context. However, backbone conformations broadly similar to the left-handed helical conformations were noted in our studies (Fig. 3). In addition, despite having an elongated double aromatic side-chain structure, Bpa could access conformational space in all four quadrants of the Ramachandran plot (Table 1). The energetic comparison of the conformations (allowed/forbidden) accessed by the NCAAs was not attempted, considering the qualitative nature and context of the study. However, each trajectory was subjected to backbone conformational clustering with a cut-off of 1.2 Å. The data presented in Figs. 4, 5, and 6 indicate that in addition to the extended sheet conformations, the given NCAAs can also induce helical or polyproline-II type conformations in the tripeptide backbone (Table 2).

Modeling of Helical peptides with the selected NCAAs

To further probe the conformational propensities of the given NCCAs in the context of the polypeptide structure,
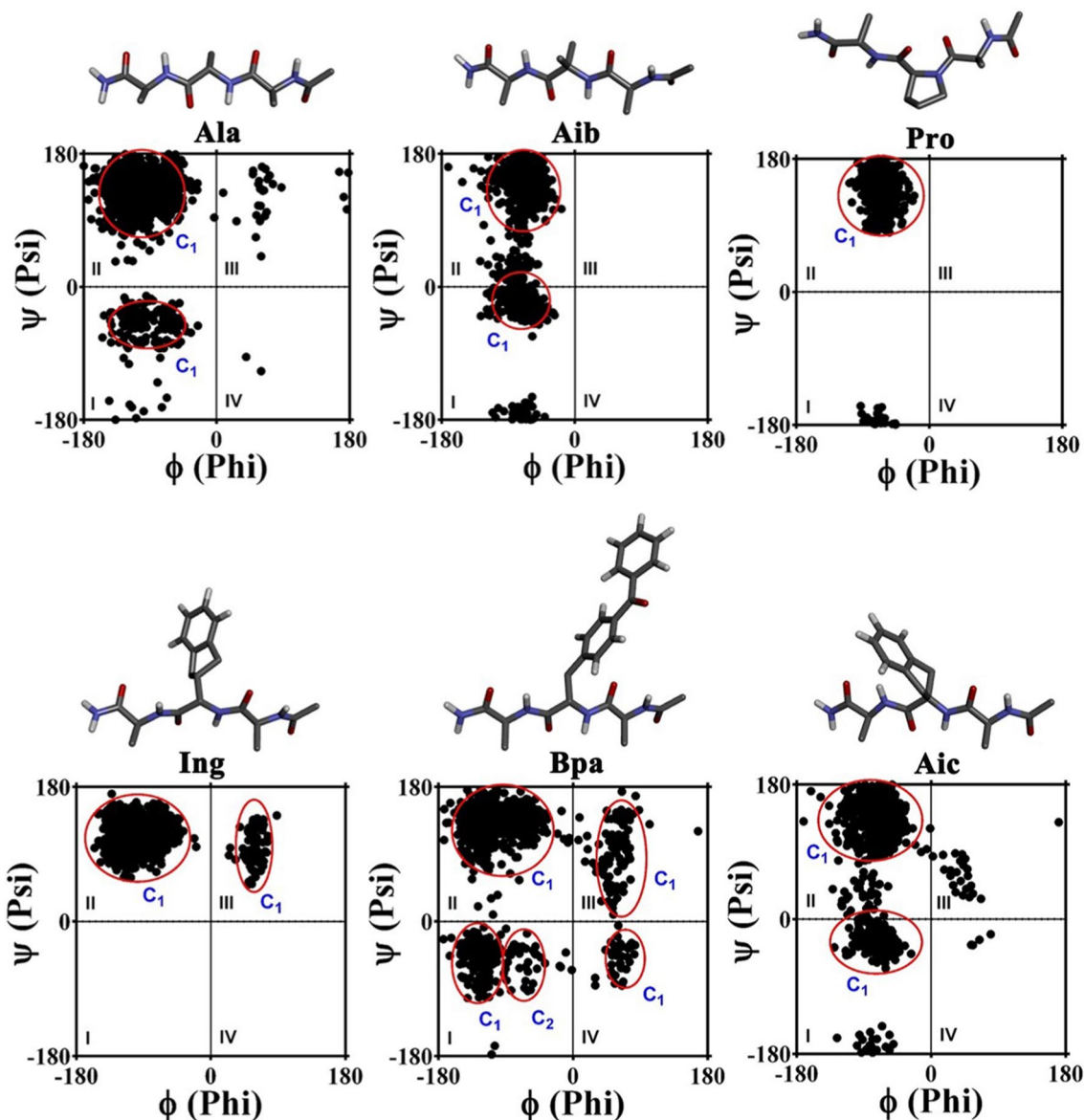
Fig. 3 Ramachandran plot illustrating the broad conformational landscapes of the NCAAs compared to the coded amino acids in the context of Ac-Ala-X-Ala-NH$_2$ tripeptide over 10 ns of MD at 300 K. The red circles highlight the most populated clusters in each quadrant of the Ramachandran plot

five 18-mer peptides (Hed1, Hed2, Hed3, Hed4, and Hed5) with almost identical sequences were modeled in the right-handed helical conformations. Among the five peptides, Hed1 [Ac-YGKAAAAK**A**AAAKAAAAK-NH$_2$] was designed to serve as the positive control. The Hed2 peptide was designed to serve as a negative control by introducing the **Ala9/Pro9** mutation in the Hed1 peptide, as proline is known to be a helix breaker under general solvent conditions. On the other hand, Hed3, Hed4, and Hed5 peptides harbored the **Ala9/Ing9, Ala9/Bpa9, and Ala9/Aic9** mutation, respectively, so the effect of these NCAAs on the helical conformation of the peptide can be effectively probed. Subsequently, the peptides were subjected to 100 ns MD studies, each in the presence of explicit water with 0.15 M NaCl at 300 K. As presented in Fig. 7, end fraying is observed across all the peptides, except the Hed2 peptide, which demonstrates a poorly folded random coil-like structure under the given conditions. Further, in addition to the Hed1, all other peptides could also sustain helical conformation to a different extent despite having the NCAAs. Overall, the data suggest that the selected NCAAs can potentially support the helical conformation to a great extent in the context of sequence-optimized polypeptides.

Table 1 The ϕ, ψ angles (mean $\pm$ SD) observed for the amino acids in the most populated clusters

| Quadrant	I	II	III	IV
Ala	C1: $\phi = -94° \pm 26°$ $\psi = -45° \pm 22°$	C1: $\phi = -100° \pm 26°$ $\psi = 131° \pm 23°$		
Aib	C1: $\phi = -66° \pm 10°$ $\psi = -26° \pm 10°$	C1: $\phi = -65° \pm 12°$ $\psi = 144° \pm 18°$		
Pro		C1: $\phi = -63° \pm 10°$ $\psi = 139° \pm 19°$		
Ing		C1: $\phi = -98° \pm 23°$ $\psi = 118° \pm 17°$	C1: $\phi = 59° \pm 7°$ $\psi = 98° \pm 23°$	
Bpa	C1: $\phi = -123° \pm 13°$ $\psi = -48° \pm 20°$ C2: $\phi = -67° \pm 14°$ $\psi = -52° \pm 20°$	C1: $\phi = -103° \pm 29°$ $\psi = 132° \pm 19°$	C1: $\phi = 57° \pm 11°$ $\psi = 94° \pm 35°$	C1: $\phi = 63° \pm 10°$ $\psi = -46° \pm 17°$
Aic	C1: $\phi = -69° \pm 13°$ $\psi = -29° \pm 10°$	C1: $\phi = -70° \pm 17°$ $\psi = 135° \pm 22°$		

Fig. 4 Schematic illustration of the central conformers of the most populated clusters evolved for the Ac-Ala-**Ing**-Ala-NH$_2$ tripeptide over the duration of MD. The data are representative of both the "endo" and "exo" trajectory. The number within the parenthesis indicates the total number of conformers for the given major cluster. The dotted green lines highlight the backbone hydrogen bonds
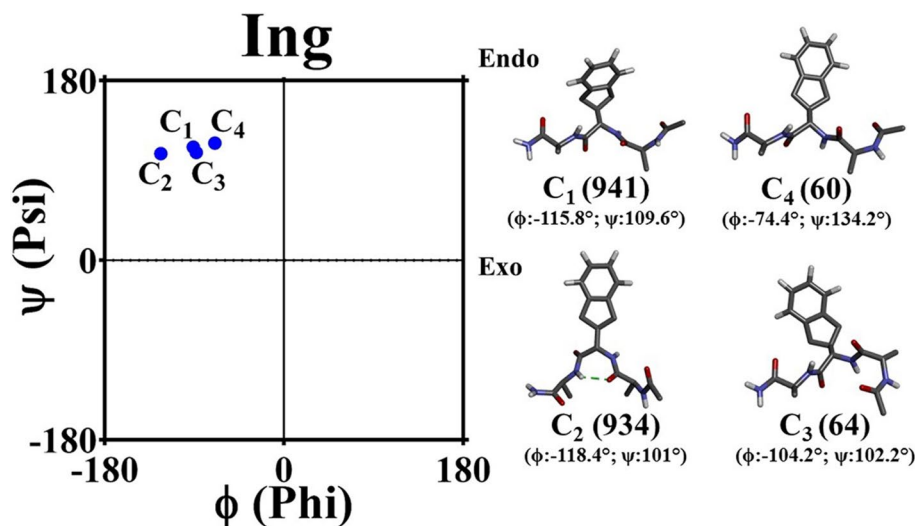
Fig. 5 Schematic illustration of the central conformers of the most populated clusters evolved for the Ac-Ala-**Bpa**-Ala-NH$_2$ tripeptide over the duration of MD. The number within the parenthesis indicates the total number of conformers for the given major cluster. The dotted green lines highlight the backbone hydrogen bonds
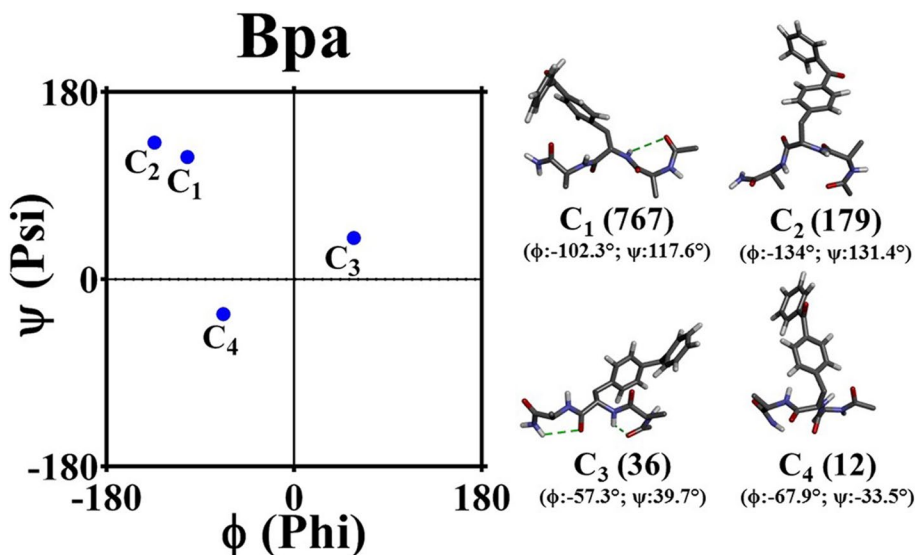
Fig. 6 Schematic illustration of the central conformers of the most populated clusters evolved for the Ac-Ala-**Aic**-Ala-NH$_2$ tripeptide over the duration of MD. The number within the parenthesis indicates the total number of conformers for the given major cluster. The dotted green lines highlight the backbone hydrogen bonds
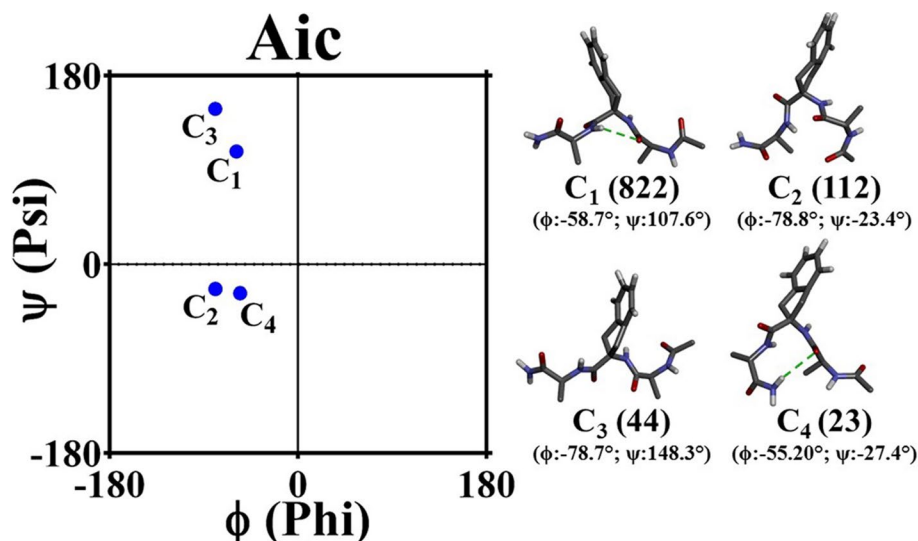
Table 2 The probable backbone conformational propensities of the selected NCAAs ascertained from the cluster analysis

Clusters	Ing	Bpa	Aic
C$_1$	Parallel β-sheet	Parallel β-sheet	β-sheet
C$_2$	Parallel β-sheet	Anti-parallel β-sheet	α$_R$-helix
C$_3$	Parallel β-sheet	α$_L$-helix	Poly-Pro helix-II
C$_4$	Poly-Pro helix-II	α$_R$-helix	3$_{10}$-helix

Preparation of the synthetic peptides

The MD data (Fig. 7) suggested that the model peptides could maintain helical conformations to a certain extent in the presence of explicit water molecules. Thus, the model peptides were synthetically prepared for experimental validation by recruiting the well-established solid-phase peptide synthesis procedures and further purified to homogeneity. Except for Hed5, all other peptides were end-capped (acylated and amidated), respectively, at the N- and C-terminus (Fig. 8). The sequence integrity of the peptides was verified by mass spectrometry (Fig. S2 to S6). The solubility check experiment indicated that the synthetic peptides are soluble in water.

It is noteworthy that acylation and amidation of the *N*-terminus and *C*-terminus of the peptides increase two amide bonds, which can not only improve the conformational stability of the short peptides but also can enhance proteolytic stability and antimicrobial potency of the synthetic peptides. In fact, NT-acylation is an established in vivo mechanism for moderating the structure, folding, and function of the proteins (Ree et al. 2018). However, several physiologically active peptides do not undergo NT-acylation, as it affects the binding of the peptides to their cognate receptors and subsequent signaling. Thus, as a control, the Hed5 peptide was not subjected to NT-acylation or CT-amidation for a better comparative understanding of the conformational effect in different solvent conditions.

Solution conformation of the synthetic peptides

The influence of the NCAAs on the overall conformation of the synthetic peptides was evaluated further by far-UV circular dichroism (CD) spectroscopy at 25 °C in the presence of various solvents (water, PBS buffer, methanol), cosolvents (TFE: trifluoroethanol, HFIP: hexafluoroisopropanol) and aqueous micellar conditions (SDS: sodium dodecyl sulfate). As presented in Fig. 9, characteristic negative CD bands, respectively at 208 and 222 nm, indicate helical conformation in the Hed1 peptide. On the other hand, the typical negative CD band around 190 nm confirms a random coil-like conformation for the Hed2 peptide in pure water (pH ~ 5.8) and 1×PBS, confirming the design principles.

It is noteworthy that in comparison to Hed1, Hed2 peptide harbors Ala9/Pro mutation, which is responsible for its random coil-like structure in water, as observed in the MD studies (Fig. 7). Interestingly, helical conformation is also observed for the Hed3, Hed4, and Hed5 peptides, mutated with Ala9/Ing, Ala9/Bpa, and Ala9/Aic NCAAs in pure water. However, the observed CD intensities for these peptides were marginally lower than the Hed1 peptide. However, all the peptides, including the Hed2 peptide, demonstrated perfect helical conformations in the presence of methanol, 20–30% TFE, 20–30% HFIP, including SDS micelles, indicating better conformational ordering, possibly due to the minimal end fraying in the model peptides under the influence of the cosolvents. The estimated percentage helicity of the peptides (Table S1) indicates 28–67% helicity for the Hed1, 0–47% helicity for the Hed2, 20–53% helicity for the Hed3, 9–14% helicity for the Hed4, and 14–43% helicity for

Fig. 7 Summary of the MD analysis data for the model peptides. The decrease in the total number of helical hydrogen bonds indicates the observed end fraying of the peptides. The central conformers of the top three most populated clusters for each model peptide are presented in cartoon representation
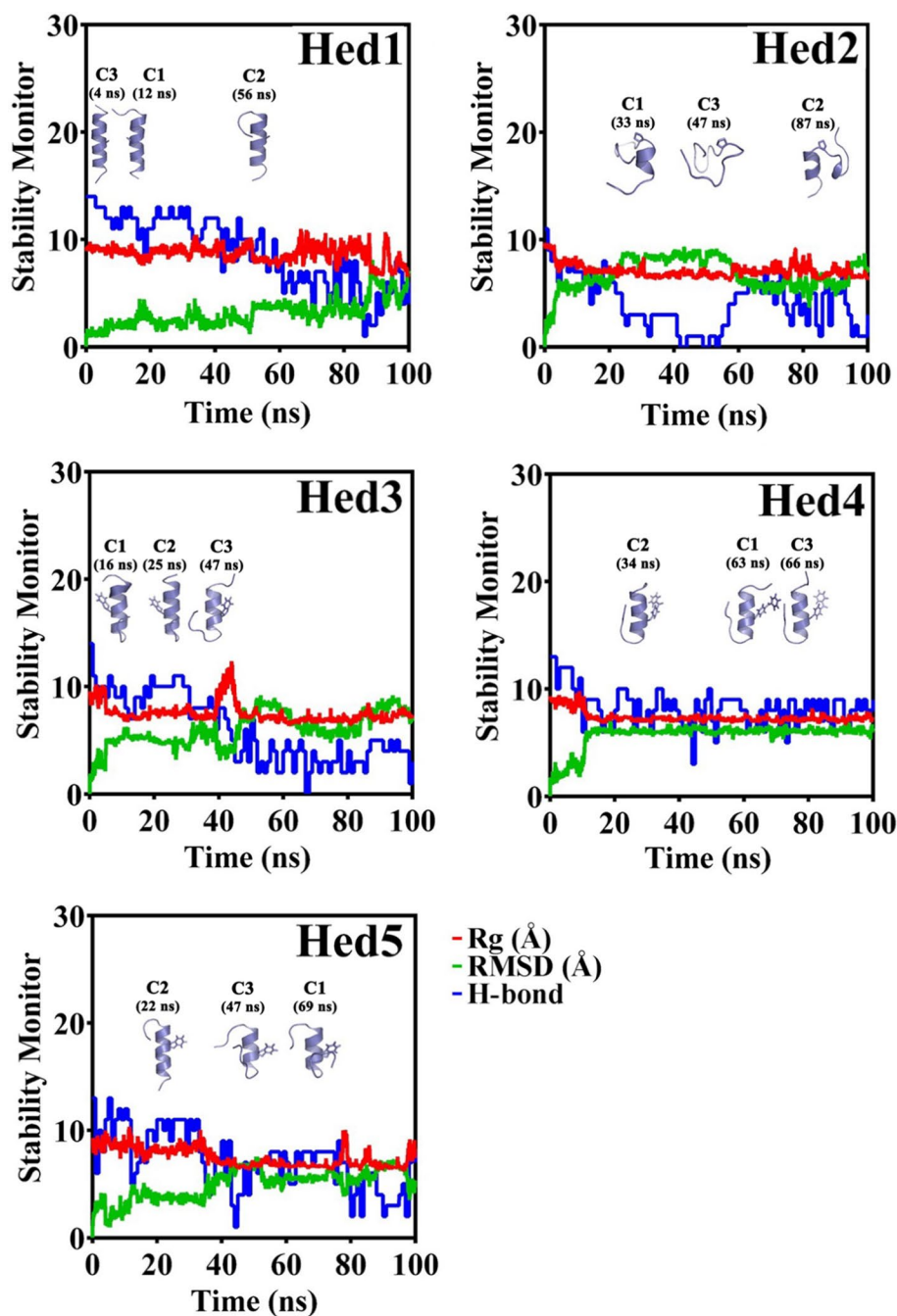
the Hed5 peptide across all the solvent conditions at 25 °C. Though the 222/208 ratio has been calculated for the peptides under all the solvent conditions, attributing the data to delineate helical character ($\alpha/3_{10}$) has not been attempted due to the simple nature of the peptide sequences. Overall the data suggest that the tested NCAAs can support ordered helical conformation in a sequence-optimized short peptides, as substitution of NCAAs did not alter the conformational state of the synthetic peptides.

Further, it was also essential to assess whether the tested NCCAs could contribute toward the thermostability of the peptides. It is well established that in addition to the amino acid sequence, secondary structure content and folding of the peptides are both solvent and temperature-dependent. However, a highly ordered folded peptide does not consistently demonstrate higher thermostability. A well-folded protein or peptide can possess a highly asymmetric secondary structural elements such as α-helix. Still, when it unfolds, it reveals loss in secondary structure, which is observable in the change in CD band intensities. Thus, recording CD spectra as a function of temperature are often helpful for evaluating the effects of mutations on the overall stability of

Peptide	Sequence	$T_{mol.\ wt.}$	$O_{mol.\ wt.}$	Purity
Hed1	Ac-YGKAAAAK-**Ala**-AAAKAAAAK-NH$_2$	1644.92	1645.2	95.1%
Hed2	Ac-YGKAAAAK-**Pro**-AAAKAAAAK-NH$_2$	1670.96	1671.2	99.0%
Hed3	Ac-YGKAAAAK-**Ing**-AAAKAAAAK-NH$_2$	1747.10	1747.2	97.8%
Hed4	Ac-YGKAAAAK-**Bpa**-AAAKAAAAK-NH$_2$	1825.13	1825.2	97.0%
Hed5	$^+_3$HN-YGKAAAAK-**Aic**-AAAKAAAAK-COO$^-$	1691.98	1691.6	95.3%

Fig. 8 Comparison of the amino acid sequences of the model synthetic peptides. The amino acids highlighted in color indicate the position-specific difference between the peptide sequences
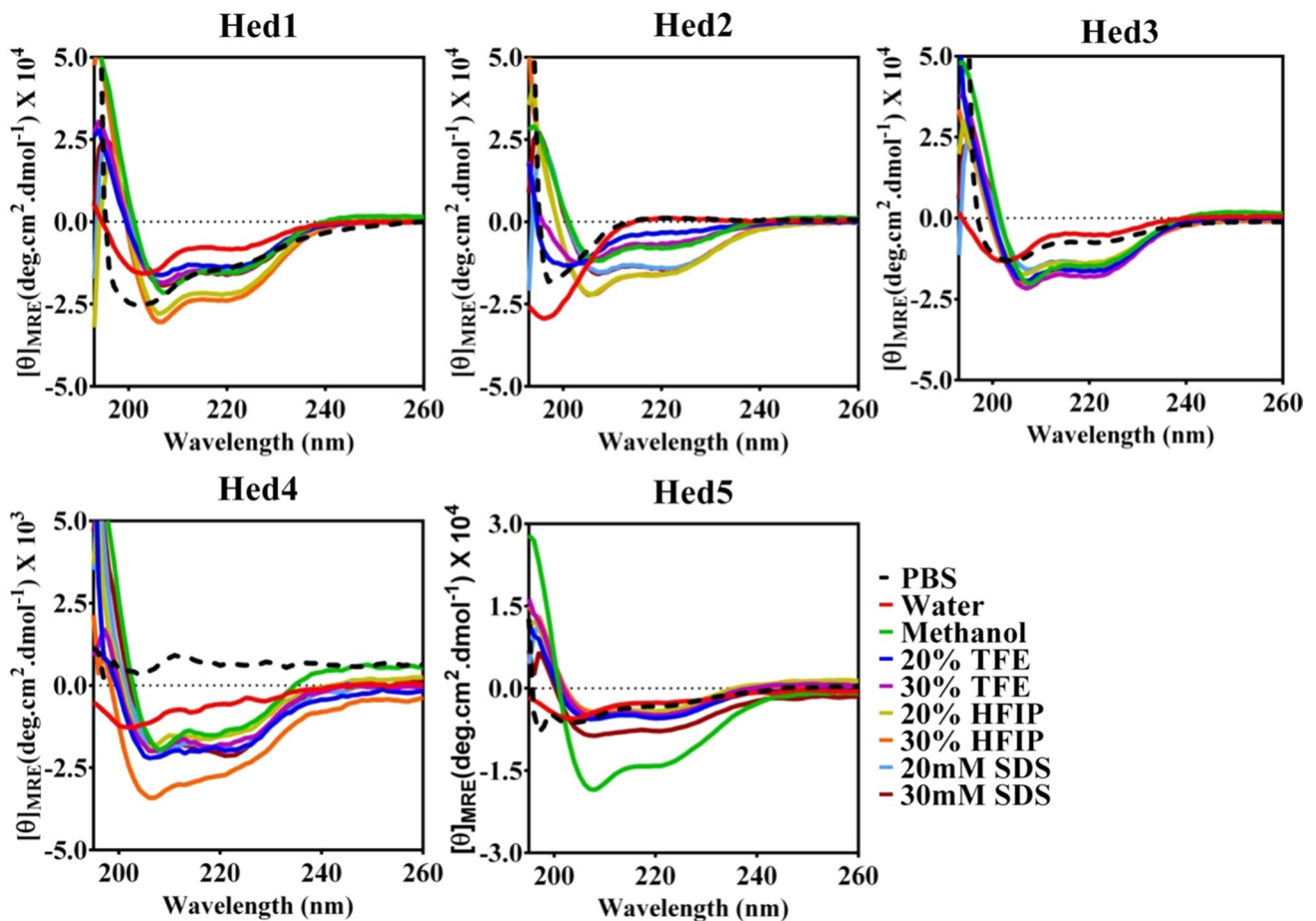

Fig. 9 Far-UV CD signature of the model synthetic peptides (50 μM) as observed under the influence of various solvents and cosolvents at 25 °C

the peptide or protein. For a fair comparison of thermostability, all the model peptides were subjected to a temperature ramping between 4–90 °C in the presence of 20% HFIP. The data presented in Fig. 10 clearly indicate that the model peptides have a well-ordered helical conformation at the low temperature, which subsequently changes with the increase in the solvent temperature. Interestingly, no other unfolding

intermediates are observed for the peptides throughout the thermal denaturation studies.

The melting temperature of the model synthetic peptides estimated from the data and the percentage of helicity estimated at the initial and the final solvent temperature are presented in Table 3. It is noteworthy that both Hed1 and Hed2 peptides having only canonical amino acids in

Fig. 10 Thermal denaturation (4–90°C) profiles of the synthetic model peptides (50 μM) recorded in 20% HFIP highlight the melting temperatures (T_m) of the respective peptides
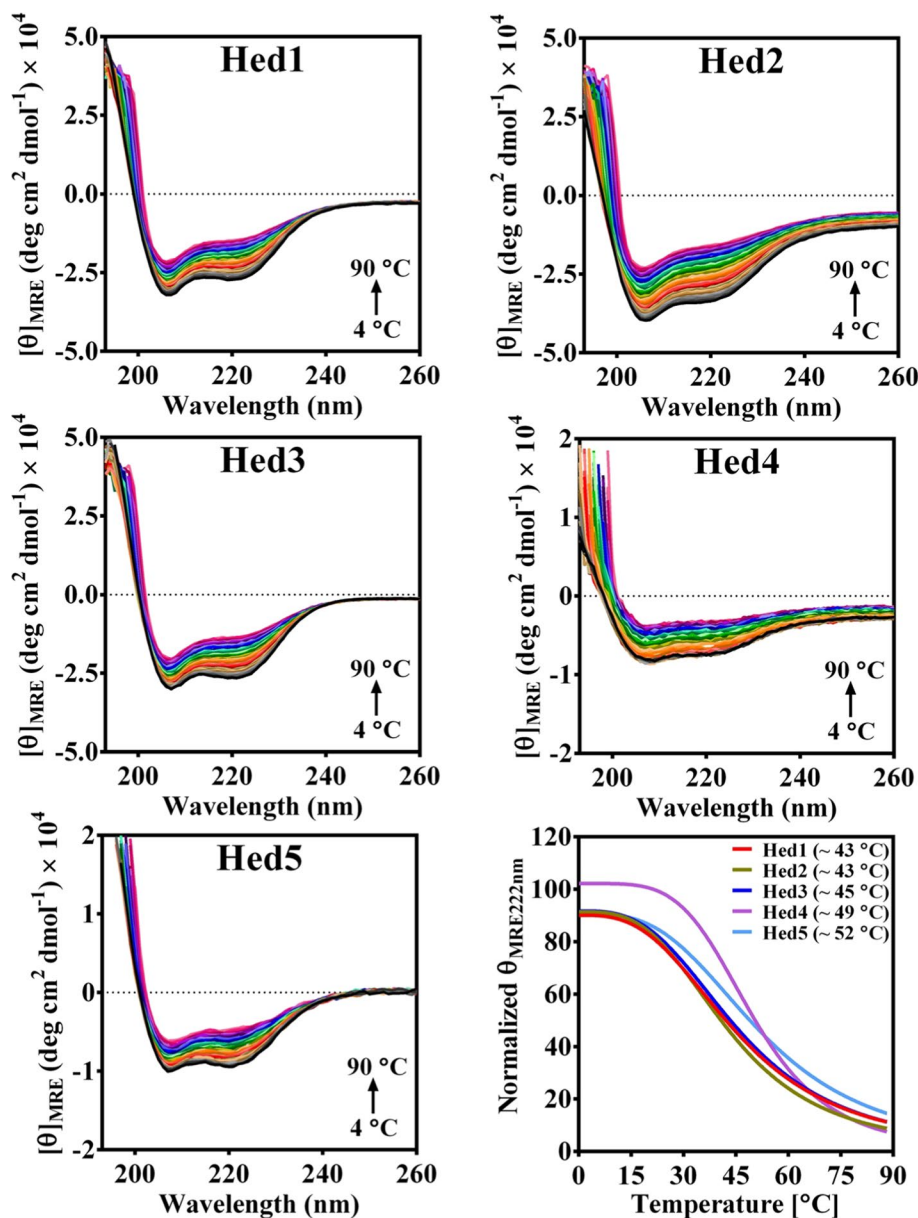
their sequences demonstrated an almost similar denaturation profile with an estimated T_m of 43 °C. In addition, the Hed3 peptide, which harbors Ala9/Ing mutation, also demonstrated a similar denaturation profile but with a slightly higher T_m of 45 °C. In contrast, Hed4 (Ala9/Bpa) and Hed5 (Ala9/Aic) peptides demonstrated a similar denaturation profile, which is different from the Hed1, Hed2, and Hed3 peptides. More importantly, both Hed4 and Hed5 peptides displayed T_m of 49 °C and 52 °C, which is significantly higher than the other peptides. Overall, the data suggest that careful usage of such NCAAs in the rational

Table 3 Melting temperature (T_m) and the percentage helicity estimated from the thermal denaturation profile of the model peptides

| Peptide	T_m (°C)	Helicity (%) at (4 °C)	Helicity (%) at (90 °C)
Hed1	43	85	45
Hed2	43	91	47
Hed3	**45**	75	42
Hed4	**49**	26	16
Hed5	**52**	32	19

design of peptide sequences can substantially improve the physical property of the peptides.

The ^{1}H-NMR of the model synthetic peptides

The CD data (Fig. 9) suggest that the synthetic model peptides acquire better conformational ordering in the presence of methanol. Thus, 1D ^{1}H-NMR of the peptides was recorded in CD$_3$OD at 25 °C. The chemical shift dispersion observed for the peptides in 1D ^{1}H-NMR data (Fig. 11) agrees with the CD data and strongly suggests the ordered nature of the peptides in the solution. The ^{1}H-NMR spectra are concentration-independent in both chemical shifts and line widths in the 0.2–2 mM concentration range in CD$_3$OD, which suggests that the peptides are free from the aggregation in the working concentration range (Fig. S7). Given the nature of the model peptide sequences, determination of the solution structure of the peptides through 2D ^{1}H-NMR was not attempted. However, an attempt was made to garner broad evidence from the observed (Fig. 8)

deviation of C$^\alpha$H chemical shift ($\Delta\delta_{H\alpha}$), which is often referred to as the chemical shift index (CSI) and used as an indicator to judge the secondary structures in the peptides and proteins (Wishart et al. 1992). The hypothetical CSI plot presented for the peptides (Fig. 12) indicates that most of the amino acids are in helical conformation, and the CD signature observed for the peptides is likely due to the continuous helical conformation extended throughout the peptide backbone and is not localized to a patch of amino acids.

In addition, as illustrated in Fig. 7, all the peptides maintain helical conformation in the water over the duration of the MD, except Hed2, which demonstrates short helical/random conformation due to the presence of a proline at the central part of the sequence. It is well known that proline acts as a helix breaker under general solvent conditions but can support helical conformation both in the presence of micelles and organic solvents (Li et al. 1996). Thus, the CSI plot presented in Fig. 12 for the peptides in CD$_3$OD depicts a realistic conformational state of
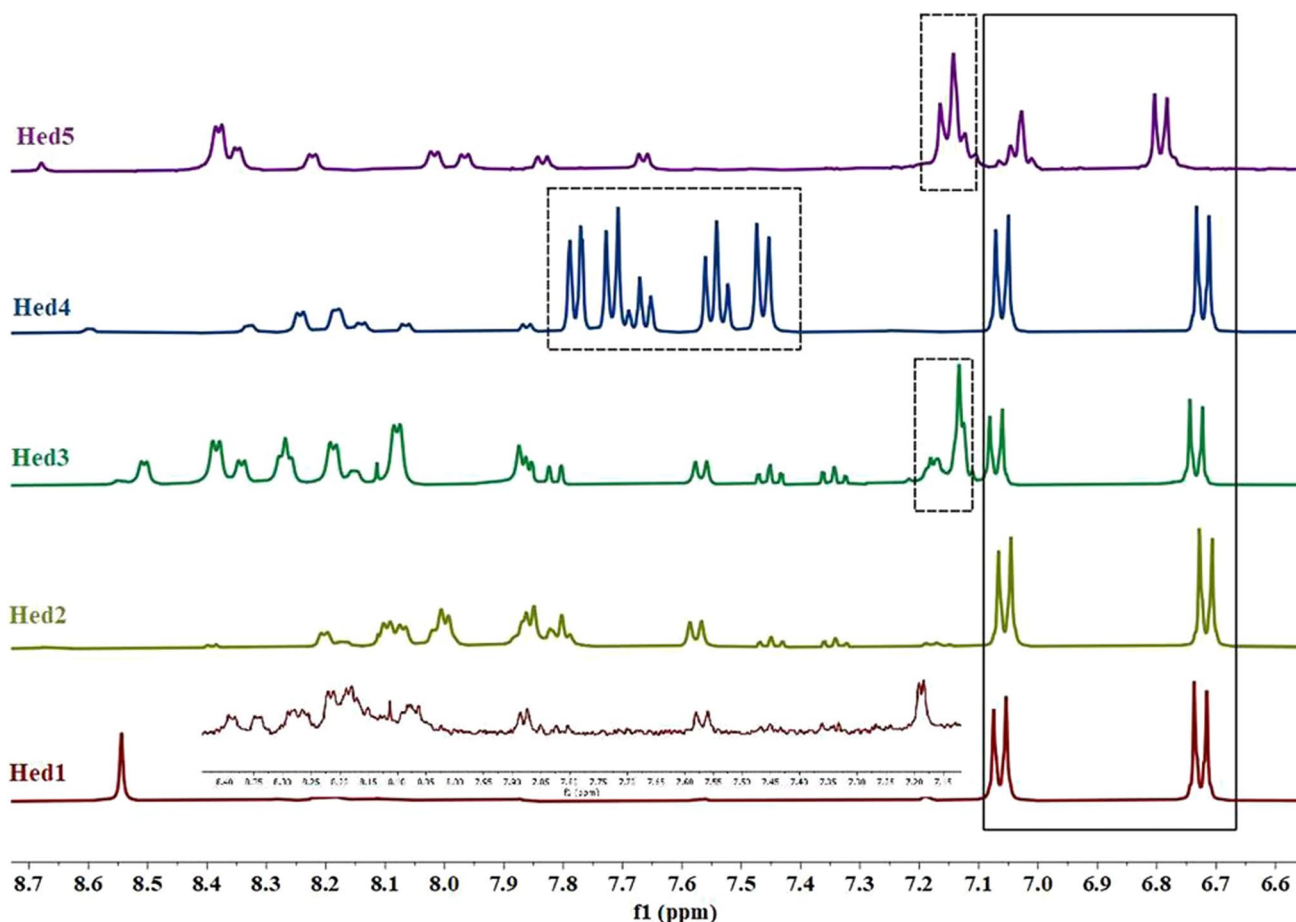
Fig. 11 The ^{1}H-NMR of the model synthetic peptides (400 MHz, ~2 mM) in CD3OD at 298 K highlights the chemical shift dispersion in the NH and aromatic regions. The tyrosine aromatic protons for the peptides are shown in the solid box, whereas the aromatic protons for Ing (Hed3), Bpa (Hed4), and Aic (Hed5) are, respectively, shown within dotted boxes

Fig. 12 Schematic illustration of the chemical shift index ($\Delta\delta_{H\alpha}$) of the amino acids of the model synthetic peptides in reference to the average random coil values observed in CD$_3$OD, suggesting the helical conformation in the peptides
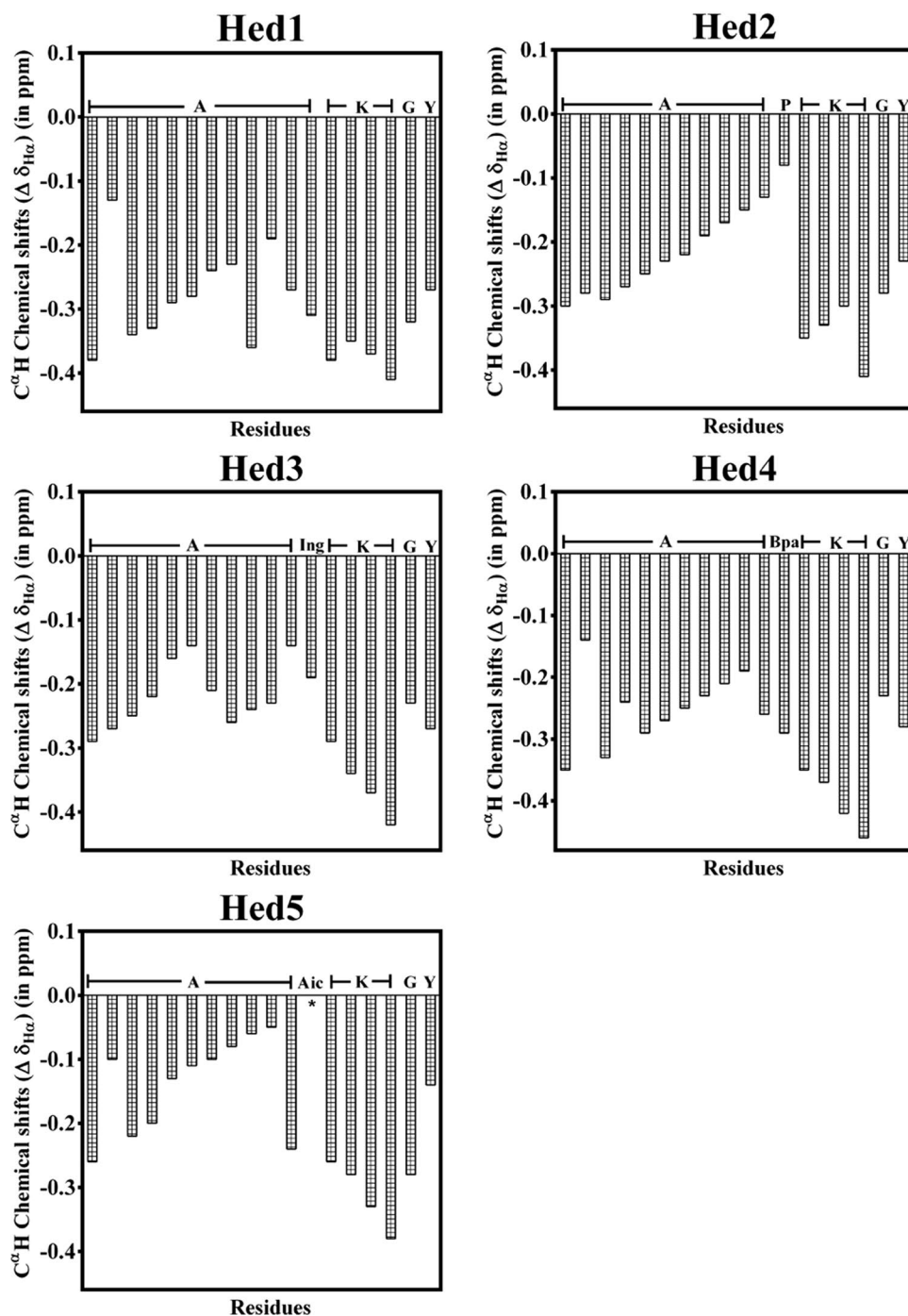
the individual amino acids contributing toward the overall helical conformation observed for the peptides in the CD studies (Fig. 9).

Fluorescence spectra of the model synthetic peptides

The model synthetic peptides have a tyrosine amino acid at the *N*-terminus. Thus, the tyrosine fluorescence was recorded

for all the peptides at 50 μM concentration in the presence of various solvents and cosolvents at 25 °C to understand the role of the tested NCCAs on the fluorescence signal of the peptides.

As noted in Fig. 13, all the peptides demonstrate an emission maximum at around 310 nm in agreement with the tyrosine fluorescence, which does not shift significantly under the influence of the different solvents and cosolvents. In agreement with the CD data, the highest fluorescence
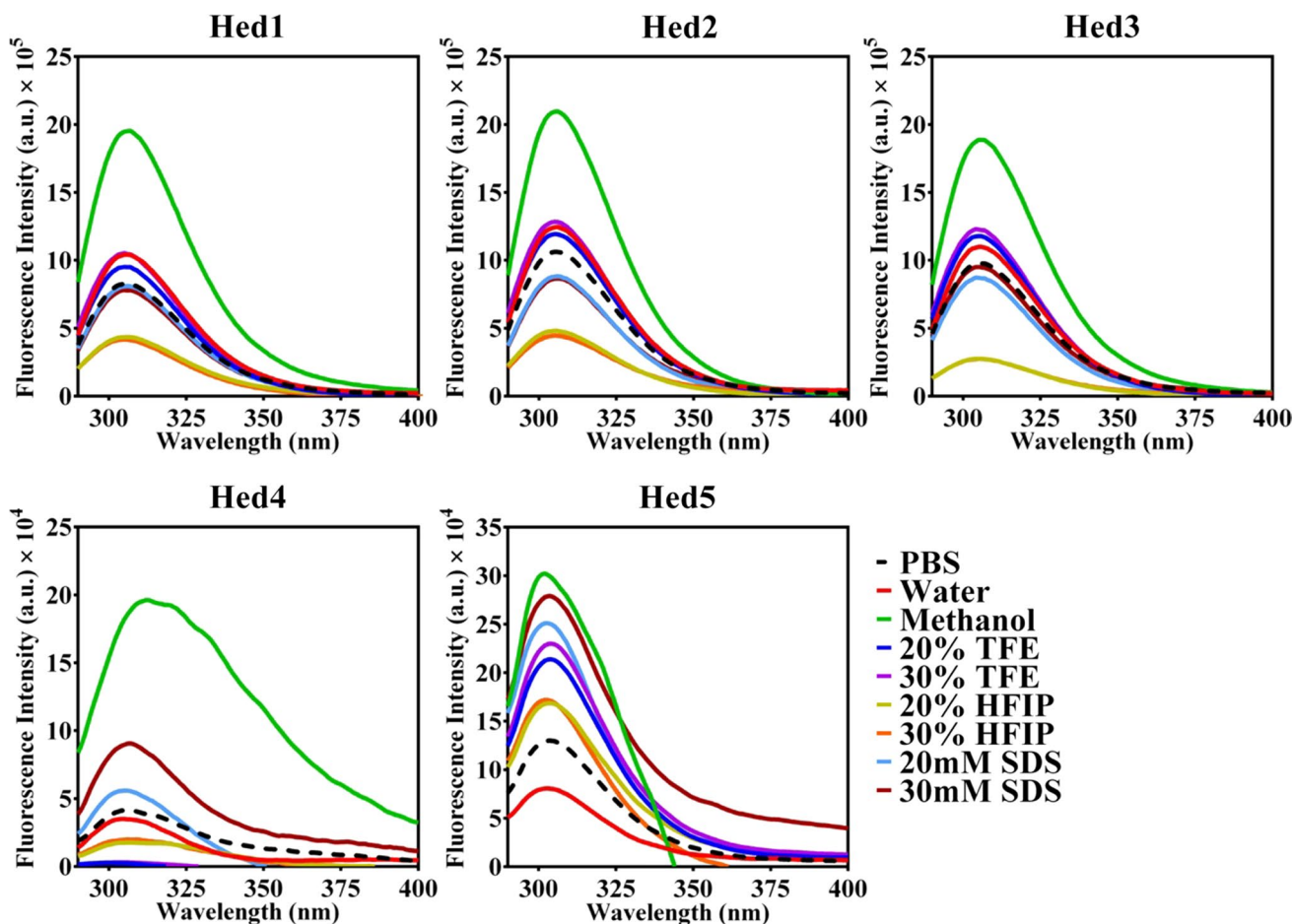
Fig. 13 Tyrosine fluorescence emission spectra of 50 μM model synthetic peptides in the presence of various solvents and cosolvents at 25 °C

intensities are also noted for the model peptides in the methanol. Despite the better conformational ordering, the fluorescence intensity of the peptides is consistently lower in the presence of TFE and HFIP compared to methanol, which indicates the solvent-induced fluorescence quenching. However, it is observed that in contrast to Hed1, Hed2, and Hed3 peptides, the fluorescence intensity of the Hed4 and Hed5 peptides in methanol is considerably lower, which could be most likely due to strong intramolecular π–π interaction (Vaiana et al. 2003), involving the aromatic side chains of Bpa and Aic respectively, with the tyrosine side chain, as observed in the MD studies (Fig. 14).

Discussion

The incorporation of the NCAAs in the rational design of peptides has the power to untap the limitless potential of the NCAAs in the discovery and engineering of peptide-based future therapeutics. Thus, it is highly pertinent to systematically explore and understand the conformational
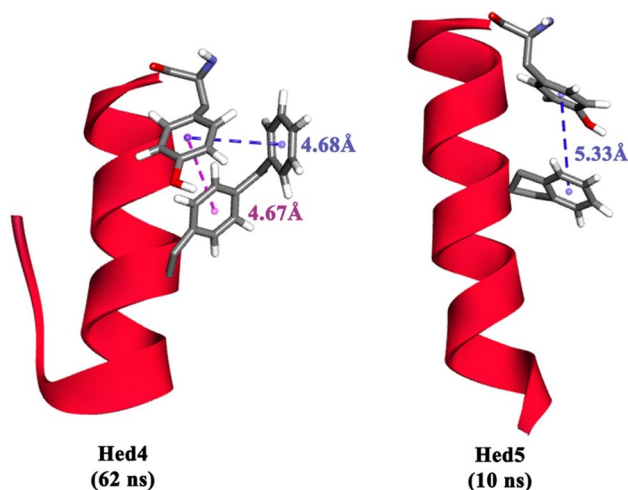

Fig. 14 Structural snapshots extracted from the MD trajectories highlight the potential intramolecular π–π interaction between the tyrosine and the aromatic side chains of Bpa and Ing, respectively, in Hed4 and Hed5 peptides

propensities of the various NCAAs in realistic model peptide sequences under a variety of solvent conditions. The current study has attempted to provide a glimpse of the possible approach in a few selected aromatic NCAAs, which can be further expanded to other types of NCAAs. The data presented for the model 18-mer peptides indicate that the selected aromatic NCAAs are well tolerated in the peptide backbone and are able to support the helical conformation of the peptides across the solvent conditions. Thus, the NCAAs can be further utilized in the rational engineering of short helices in sequence-optimized peptides. Nevertheless, the potential of the selected NCAAs to influence β-sheet conformation needs to be further probed in a sequence and structure-dependent manner. In addition to the conformational ordering (Figs. 7, 9, and 12), it is also observed that the selected NCAAs can enhance the thermal stability of the model peptides (Fig. 10) by several degrees. Peptides with NCAAs are generally proteolytically resistant, and enhancement of thermostability of such peptides is a desired pharmacokinetic gain required for the development of peptide-based therapeutics. Further, it is noteworthy that the selected aromatic NCAAs can forge both hydrophobic as well as other types of inter and intramolecular interactions, such as "π–π", "cation–π", "sulfur–π", "alkyl–π", including unconventional weak hydrogen bonding with the neighboring amino acids,

which can enhance the overall molecular recognition ability of the designer peptides for targeting physiologically relevant proteins, enzymes and receptors. Moreover, combining the power of computational biology, the model 18-mer peptides described in the current study can be further optimized in terms of sequence, structure, and overall length (Joshi et al. 2006) by utilizing a proper mix of both coded and all kinds of NCAAs. Recruitment of the NCAAs can not only improve the binding affinity and stability but also can significantly enhance the druggability and formulation capabilities of the synthetic peptides as future drugs, which can be exploited in a variety of therapeutic applications as illustrated schematically in Fig. 15.

For example, the glucagon-like peptide-binding receptor (GLP-1R) is an established drug target for the management of type-II diabetes mellitus (T2DM). GLP-1R binds to GLP-1 peptide, an incretin hormone made of 30 or 31-amino acids, and triggers the secretion of insulin from the pancreatic β-cells in a glucose-dependent manner. However, GLP-1 is quickly metabolized by the enzyme dipeptidyl peptidase-4 (DPP4), resulting in the inactivation of the GLP-1. The usage of NCAA like Aib has been useful in the successful design of 30-mer peptides, like semaglutide and taspoglutide, which are being marketed as FDA-approved peptide drugs for the management of T2DM. The recent discovery and translation of small molecule drugs targeting GLP-1R to
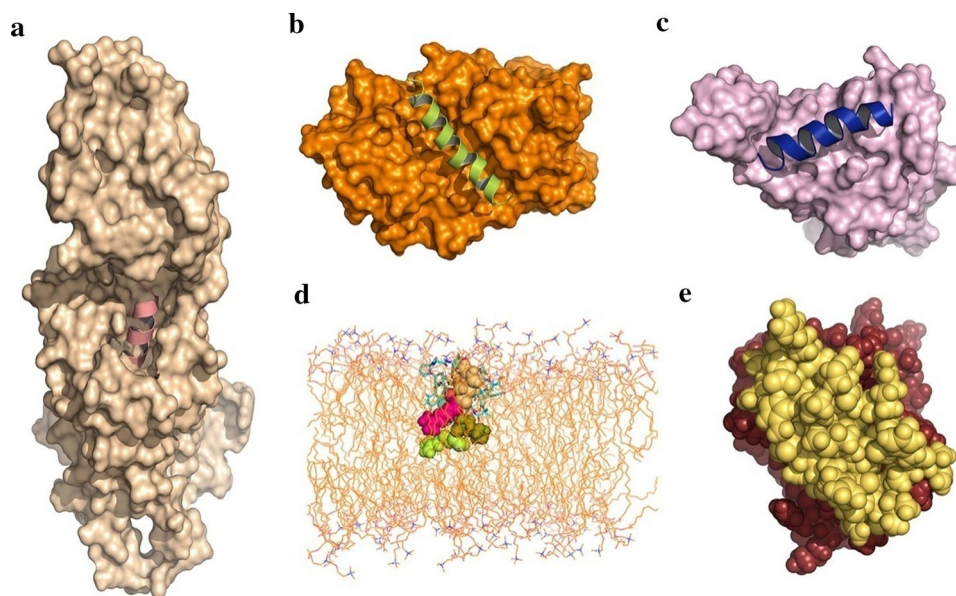
Fig. 15 Potential therapeutic applications of the sequence-optimized designer peptides ≤ 30 amino acids composed of both coded and NCAAs. **a** Targeting the GLP-1R (wheat, PDB: 6XOX) through short designer peptide agonists (GLP-1 mimetic, salmon pink) for the management of T2DM, **b** Targeting the PPIs involving the G-protein subunits (G$_\alpha$ subunit, orange, PDB: 2ODE)] and the RGS protein through a designer peptide (pea green) inhibitor, **c** Targeting the PPIs involving the ACE2 receptor and the receptor-binding domain of the SARS-CoV2 spike protein (light pink, PDB: 6M0J) through designer small peptide inhibitors (blue), **d** Illustration of the membrane disruption potential of a designer small antimicrobial peptide (CPK) in the presence of a model POPC bilayer with a focus on the antimicrobial resistance (AMR), and **e** Design of antibody-like neutralizing peptide (yellow) targeting the complement protein C5a (brick red, PDB: 1KJS) for the treatment of chronic inflammation-induced diseases

clinical trials is very encouraging (Cong et al. 2022; Kawai et al. 2020), which suggests that conformationally ordered shorter peptide agonists with better pharmacokinetic properties can be designed for targeting the GLP-1R (Fig. 15a).

Similarly, designer peptides, both optimized in sequence and size, can be utilized as competitive inhibitors targeting pharmacologically important PPIs. For example, downstream signaling of GPCRs can be modulated alternatively by targeting the regulator of G-protein signaling (RGS) proteins (Soundararajan et al. 2008), which upon binding, enhances the intrinsic GTPase activity of the Gα-subunit of heterotrimeric G-protein and thus, control the duration of the agonist-induced signaling cycle of the GPCRs. Targeting the PPIs involving the G-protein and RGS protein through designer peptides (Fig. 15b) deliverable to the cytosolic environment of the cells has the potential to reduce the therapeutic dose size and the off-target actions of certain GPCR agonists. In addition, such peptides can also enhance the natural signaling output of the GPCRs challenged by the poor level of natural agonists under certain disease conditions like neuronal disorder and depression (Squires et al. 2018). A similar approach can also be followed for blocking the PPIs (Fig. 15c) of the SARS-CoV2 spike protein (Lan et al. 2020) with the host cell angiotensin-converting enzyme-2 (ACE2) receptor for minimizing the onset of severe COVID19 in desired populations. It is noteworthy that most of the peptide mimetics reported for blocking the interaction of ACE2 (Pomplun 2020) with the spike protein of the SARS-CoV2 are in the range of 25–65 amino acids, and many of them do not harbor NCAAs. Thus, there is a huge potential for exploring sequence-optimized conformationally ordered shorter peptides involving NCAAs for therapeutic applications.

Besides planning to control the current pandemic, it is also essential to stay prepared for the future health problems that are going to surface in the post-COVID19 world. Antimicrobial resistance (AMR) may aggravate further both during the pandemic and also beyond the pandemic, a major concern that requires urgent attention (Dadgostar 2019; Upert et al. 2021), as there is a dearth of new-age antimicrobials (AMP) in the developmental pipeline. Further, AMPs are preferred (Magana et al. 2020) over conventional antibiotics, as they can modulate the immune response appropriately, and they have better antibiofilm activity, including slower emergence of resistance. The perceived resistance to proteolysis, thermostability, and the ability to induce the desired conformation in short peptides suggest that careful inclusion of the NCAAs into the designer peptides (Fig. 15d) optimized in sequence and size can provide the required alternatives to the naturally secreted antimicrobial peptides (Huan et al. 2020) with low cytotoxicity and hemolytic activity both in vitro and in vivo settings. In fact, AMPs are an essential part of the immune system,

which are involved in controlling the signaling activities of proinflammatory cytokines in the eukaryotes. It is noteworthy that in response to stress, injury, and exposure to the contagions like the common influenza viruses (H5N1, H7N9), and the severe acute respiratory syndrome (SARS) coronavirus, an overwhelmed immune system can aberrantly upregulate the concentration of the complement fragment 5a (C5a), (Zhang et al. 1997) an extremely potent proinflammatory glycoprotein of the complement system, augmenting the lethal "cytokine storm" that exacerbate several chronic inflammations induced pathological conditions, such as sepsis, allergic asthma, acute lung injury, including the moderate to severe COVID19 in humans. Indeed, C5a has been found to be significantly elevated to several folds in the plasma of moderate to severe COVID19 patients (Carvelli et al. 2020) compared to the healthy subjects. Thus, C5a is considered to be a major pharmacological target, and no wonder that ~ 20 drug candidates targeting the various stages of the complement signaling are currently under development by several pharma majors, and most of the drugs in the process of development are antibodies. In fact, Eculizumab (Soliris; Alexion Pharmaceuticals) and IFX-1 (InflaRx) are the two FDA-approved antibodies currently available in the market that respectively targets C5, the precursor protein of C5a, or the C5a directly. C5a being an important protein of the immune system, completely shutting it down through antibodies is physiologically undesired, and thus, we have been actively pursuing the avenue of repurposing the FDA-approved drugs as potential "neutraligands" of C5a (Das and Rana 2021; Mishra and Rana 2019; Mishra et al. 2022) for therapeutically modulating the pathological stimulation of C5a-C5aR1 signaling axes. Nevertheless, antisense peptides (Fujita et al. 2004) targeting the C5a have been explored earlier, and thus, in the context of the PPIs involving C5a and C5aR1 (Das et al. 2021; Sahoo et al. 2018; Rana and Sahoo 2015), the excessive C5a produced in the body can also be neutralized with designer peptides (Fig. 15e), mimicking the function of the antibodies for maintaining the desired balance between the pharmacological and physiological concentration of C5a in the plasma for the therapeutic benefits. The examples illustrated in Fig. 15 are the snapshots of the currently ongoing studies involving small designer peptides of unusual architecture involving NCAAs, which are likely to be elaborated further in subsequent future studies through experimental validation in appropriate biological systems.

Conclusions

Peptides, the molecular fragments of the proteins, are considered the ideal biological probes for understanding the structure–function relationships behind the complex protein–protein interactions of several biological processes.

The chemical versatility available to peptides makes them one of the go-to scaffolds for the rational design of targeted therapeutics involving both extracellular, transmembrane, and intracellular proteins. From the humble beginning of the synthesis of glycylglycine dipeptide to the marketing of the synthetic peptide lypressin as antidiuretics, peptide chemistry has rapidly advanced so far over the last 50 years (Sachdeva 2016). Though the synthesis of peptides has come a long way, peptides as therapeutics have yet to capture a sizable market share compared to other therapeutics. Focus on the native or the analogs of the native peptide only could be one of the possible reasons behind this. Systematic expansion of the chemical space in peptides within the boundary of conformational space is possible with de novo design of small heterologous peptides, involving a strategic mix of both coded and NCAAs, which can be exploited further for therapeutic applications. The above approach is exemplified in the current study with a set of selected C_α-substituted aromatic NCAAs in a model helical system, which can be easily expanded to any other systems by systematically screening the repertoire of synthetically available NCAAs for subsequent therapeutic explorations.

Materials and methods

Peptide modeling and Molecular dynamics (MD) simulation

The Ala-Gly-Ala tripeptide end-capped both at the *N*-terminus (-COCH3) and the *C*-terminus (-NH2) were initially modeled in an extended conformation ($\phi = -120°$, $\psi = 120°$) using the Discovery studio (Biovia). Further, the structure of Bpa, Ing, and Aic was geometrically optimized in the Discovery studio and subsequently incorporated into the Ac-Ala-Gly-Ala-NH2 tripeptide template to respectively prepare the Ac-Ala-Bpa-Ala-NH2, Ac-Ala-Ing-Ala-NH2, and Ac-Ala-Aic-Ala-NH2 tripeptides. The Ac-Ala-Ala-Ala-NH2, Ac-Ala-Aib-Ala-NH2, and Ac-Ala-Pro-Ala-NH2 tripeptides were also designed similarly to serve as the reference. The tripeptides were energy minimized in a cubic box under periodic boundary conditions both in the presence of vacuum and explicit simple point charge (SPC) water with density set to the value corresponding to 1 atm. Subsequently, the system was subjected to equilibration for 100 ps each at 300 K under NVT and NPT conditions prior to initiating the production molecular dynamics (MD) run over 10 ns by recruiting the GROMACS software package (Hess et al. 2008). The NCAAs were incorporated into the GROMACS force field by involving a fragment built-up approach to make the system work. Subsequently, a template 18-mer synthetic peptide code named as "Hed1" was modeled as an end-capped (N-terminus: -COCH3; C-terminus:-NH2)

α-helix ($\phi = -47°$, $\psi = -57°$) using the Discovery studio, by relying on the reported CD spectroscopy data. The "Hed1" peptide was further used as a template for generating the other (Hed2, Hed3, Hed4, and Hed5) model helices harboring one of the NCCAs at a strategic location. As described above, all the model helices were subjected to comparative MD studies over 100 ns each at 300 K. The MD trajectories were thoroughly analyzed as desired using the modules built into GROMACS. Data were plotted in GraphPad Prism (version 9 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com). Both PyMOL (The PyMOL Molecular Graphics System, Version 2.5.2, Schrödinger, LLC) and Discovery studio (BIOVIA) software were used to process, visualize, analyze, and present the peptide structures.

Peptide synthesis

Using the standard Fmoc chemistry over the solid phase, the designed 18-mer helical peptides (Hed1, Hed2, Hed3, Hed4, and Hed5) were synthetically prepared by recruiting the services of GenScript (NJ, USA). Except for Hed5, all peptides were acylated at the *N*-terminus and amidated at the *C*-terminus. Post cleavage, the peptides were purified, and the absolute purity ($\geq 95\%$) of the peptides was judged from the analytic HPLC profile of the peptides recorded by recruiting the C18 (4.6×250 mm) column at 220 nm, using acetonitrile–water gradient in the presence of 0.05–0.065% trifluoroacetic acid (TFA). The ESI–MS further confirmed the identity of the peptides. All the peptides were soluble in ultrapure water and 1×PBS (pH 7.4).

Circular dichroism (CD) studies

The far-UV CD spectra of the peptides were recorded by the Chirascan CD spectrometer system (Applied photophysics) at 25 °C. Each sample was subjected to a minimum of three scans with a time constant of 1 s and a step size of 1 nm. The PBS buffer and Milli Q water were filtered through a 0.2 µm filter and thoroughly degassed with continuous stirring while bubbling with nitrogen gas prior to the experiment. The stock concentration of the peptides was determined using a UV spectrophotometer. The CD spectra of the peptides were recorded at a concentration of 50 µM in water (pH ~ 5.8) and methanol, including a varying percentage of cosolvents, such as trifluoroethanol (TFE), hexafluoroisopropanol (HFIP), as well as in 20–30 mM SDS micelles suspended in PBS. The CD spectra were background corrected by subtracting the blank spectra of the corresponding solvents used in the study. The thermostability of the model peptides was judged by recording the CD spectra between 4 and 90 °C in the presence of 20% HFIP with a stepping gradient of 2 °C. The sample was equilibrated for 1 min at

each temperature point before recording the ellipticity. The CD signal in millidegrees was converted to mean residue ellipticity [θ_{MRE}deg.cm^2.dmol^{-1}] using the following equation (millidegrees × mean residual weight)/ (pathlength in millimeter × concentration in mg ml^{-1}). The mean residual weight of a peptide is the total molecular weight divided by the total number of amides in the backbone (Greenfield 2006). All data points were plotted in GraphPad Prism (version 9), and the spectra were smoothened up to 2–3 units wherever required. The helicity percentage was calculated from [θ]$_{MRE}$ at 222 nm (Morrisett et al. 1973).

Fluorescence studies

The fluorescence emission spectra of the peptides were recorded at 50 µM concentration in the presence of various solvents and cosolvents at 25 °C using the Horiba Fluoromax spectrophotometer. The excitation and emission slit widths were set to 5 nm, with excitation wavelength set at 280 nm and emission range set between 290 and 500 nm. The fluorescence spectra presented represent an average of three scans that have been corrected for the background spectra of the respective solvents.

Nuclear magnetic resonance (NMR) studies

The ^{1}H-NMR of the peptides maintained at 2 mM were recorded using a Bruker AVANCE 400 MHz NMR spectrometer equipped with cryoprobe at 298 K in 99.8% CD$_3$OD having 0.03% v/v tetramethylsilane (TMS) as the internal reference. Each peptide was subjected to a minimum of 32 scans. In addition, the ^{1}H-NMR of the peptides was also recorded at tenfold dilution to judge the molecular aggregation of the peptides. The raw FIDs were processed with the TOPSPIN provided by Bruker, followed by MestreNova software for presentation. Hypothetical C$^\alpha$H chemical shift index ($\Delta\delta_{H\alpha}$) for the peptides were generated using the relation $\Delta\delta = \delta_{obs} - \delta_{coil}$ using mean δ_{coil} values for the given residues reported in GGXGG peptide under identical solvent conditions. (Merutka et al. 1995).

Declarations

References

Ambrogelly A, Palioura S, Soll D (2007) Natural expansion of the genetic code. Nat Chem Biol 3(1):29–35. https://doi.org/10.1038/nchembio847

Antosova Z, Mackova M, Kral V, Macek T (2009) Therapeutic application of peptides and proteins: parenteral forever? Trends Biotechnol 27(11):628–635. https://doi.org/10.1016/j.tibtech.2009.07.009

Avan I, Hall CD, Katritzky AR (2014) Peptidomimetics via modifications of amino acids and peptide bonds. Chem Soc Rev 43(10):3575–3594. https://doi.org/10.1039/c3cs60384a

Balaram P (1992) Non-standard amino acids in peptide design and protein engineering. Curr Opin Struct Biol 2(6):845–851

Beck DA, Alonso DO, Inoyama D, Daggett V (2008) The intrinsic conformational propensities of the 20 naturally occurring amino acids and reflection of these propensities in proteins. Proc Natl Acad Sci U S A 105(34):12259–12264. https://doi.org/10.1073/pnas.0706527105

Bell EA (2003) Nonprotein amino acids of plants: significance in medicine, nutrition, and agriculture. J Agric Food Chem 51(10):2854–2865. https://doi.org/10.1021/jf020880w

Bell EA, Watson AA, Nash RJ (2008) Non-protein amino acids: a review of the biosynthesis and taxonomic significance. Nat Prod Commun 3(1):93–110

Bisetty K, Catalan JG, Kruger HG, Perez JJ (2005) Conformational analysis of small peptides of the type Ac–X–NHMe, where X = Gly, Ala, Aib and Cage. J Mol Struct (thoechem) 731(1–3):127–137. https://doi.org/10.1016/j.theochem.2005.04.037

Blaskovich MA (2016) Unusual Amino Acids in Medicinal Chemistry. J Med Chem 59(24):10807–10836. https://doi.org/10.1021/acs.jmedchem.6b00319

Brunel FM, Liu F, Mayer JP (2019) Trends in peptide therapeutics. Successful Drug Discovery. https://doi.org/10.1002/9783527814695.ch1

Carvelli J, Demaria O, Vely F, Batista L, Chouaki Benmansour N, Fares J, Carpentier S, Thibult ML, Morel A, Remark R, Andre P, Represa A, Piperoglou C, Explore C-IPHg, Explore C-MIg, Cordier PY, Le Dault E, Guervilly C, Simeone P, Gainnier M, Morel Y, Ebbo M, Schleinitz N, Vivier E (2020) Association of COVID-19 inflammation with activation of the C5a–C5aR1 axis. Nature 588(7836):146–150. https://doi.org/10.1038/s41586-020-2600-6

Cavaco M, Castanho MA, Neves V (2018) Peptibodies: An elegant solution for a long-standing problem. Pept Sci 110(1):e23095

Chin JW, Martin AB, King DS, Wang L, Schultz PG (2002) Addition of a photocrosslinking amino acid to the genetic code of Escherichiacoli. Proc Natl Acad Sci U S A 99(17):11020–11024. https://doi.org/10.1073/pnas.172226299

Clerici F, Erba E, Gelmi ML, Pellegrino S (2016) Non-standard amino acids and peptides: From self-assembly to nanomaterials. Tetrahedron Lett 57:5540–5550

Cong Z, Zhou Q, Li Y, Chen LN, Zhang ZC, Liang A, Liu Q, Wu X, Dai A, Xia T, Wu W, Zhang Y, Yang D, Wang MW (2022) Structural basis of peptidomimetic agonism revealed by

small- molecule GLP-1R agonists Boc5 and WB4-24. Proc Natl Acad Sci U S A 119(20):e2200155119. https://doi.org/10.1073/pnas.2200155119

Croitoru A, Park SJ, Kumar A, Lee J, Im W, MacKerell AD Jr, Aleksandrov A (2021) Additive CHARMM36 force field for non-standard amino acids. J Chem Theory Comput 17(6):3554–3570. https://doi.org/10.1021/acs.jctc.1c00254

Dadgostar P (2019) Antimicrobial resistance: implications and costs. Infect Drug Resist 12:3903–3910. https://doi.org/10.2147/IDR.S234610

Das A, Behera LM, Rana S (2021) Interaction of human C5a with the major peptide fragments of C5aR1: direct evidence in support of "two-site" binding paradigm. ACS Omega 6(35):22876–22887. https://doi.org/10.1021/acsomega.1c03400

Das A, Rana S (2021) The role of human C5a as a non-genomic target in corticosteroid therapy for management of severe COVID19. Comput Biol Chem 92:107482. https://doi.org/10.1016/j.compbiolchem.2021.107482

Davenport AP, Scully CCG, de Graaf C, Brown AJH, Maguire JJ (2020) Advances in therapeutic peptides targeting G protein-coupled receptors. Nat Rev Drug Discov 19(6):389–413. https://doi.org/10.1038/s41573-020-0062-z

Di L (2015) Strategic approaches to optimizing peptide ADME properties. AAPS J 17(1):134–143. https://doi.org/10.1208/s12248-014-9687-3

Ding Y, Ting JP, Liu J, Al-Azzam S, Pandya P, Afshar S (2020) Impact of non-proteinogenic amino acids in the discovery and development of peptide therapeutics. Amino Acids 52(9):1207–1226. https://doi.org/10.1007/s00726-020-02890-9

Durani S (2008) Protein design with L- and D-alpha-amino acid structures as the alphabet. Acc Chem Res 41(10):1301–1308. https://doi.org/10.1021/ar700265t

Fosgerau K, Hoffmann T (2015) Peptide therapeutics: current status and future directions. Drug Discov Today 20(1):122–128. https://doi.org/10.1016/j.drudis.2014.10.003

Fujita E, Farkas I, Campbell W, Baranyi L, Okada H, Okada N (2004) Inactivation of C5a anaphylatoxin by a peptide that is complementary to a region of C5a. J Immunol 172(10):6382–6387. https://doi.org/10.4049/jimmunol.172.10.6382

Gellman SH (1998) Foldamers: a manifesto. Acc Chem Res 31(4):173–180

Goodman CM, Choi S, Shandler S, DeGrado WF (2007) Foldamers as versatile frameworks for the design and evolution of function. Nat Chem Biol 3(5):252–262. https://doi.org/10.1038/nchembio876

Gray DA, Wenzel M (2020) More than a pore: a current perspective on the in vivo mode of action of the lipopeptide antibiotic daptomycin. Antibiotics (basel) 9(1):17. https://doi.org/10.3390/antibiotics9010017

Greenfield NJ (2006) Using circular dichroism spectra to estimate protein secondary structure. Nat Protoc 1(6):2876–2890. https://doi.org/10.1038/nprot.2006.202

Grunbeck A, Huber T, Sachdev P, Sakmar TP (2011) Mapping the ligand-binding site on a G protein-coupled receptor (GPCR) using genetically encoded photocrosslinkers. Biochemistry 50(17):3411–3413. https://doi.org/10.1021/bi200214r

Gupta M, Chauhan VS (2011) De novo design of alpha, beta-didehydrophenylalanine containing peptides: from models to applications. Biopolymers 95(3):161–173. https://doi.org/10.1002/bip.21561

Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 4(3):435–447. https://doi.org/10.1021/ct700301q

Huan Y, Kong Q, Mou H, Yi H (2020) Antimicrobial peptides: classification, design, application and research progress in multiple fields. Front Microbiol 11:582779. https://doi.org/10.3389/fmicb.2020.582779

Johnson JA, Lu YY, Van Deventer JA, Tirrell DA (2010) Residue-specific incorporation of non-canonical amino acids into proteins: recent developments and applications. Curr Opin Chem Biol 14(6):774–780. https://doi.org/10.1016/j.cbpa.2010.09.013

Joiner CM, Breen ME, Mapp AK (2019) Electron-deficient p-benzoyl-l-phenylalanine derivatives increase covalent chemical capture yields for protein-protein interactions. Protein Sci 28(6):1163–1170. https://doi.org/10.1002/pro.3621

Joshi S, Rana S, Wangikar P, Durani S (2006) Computational design of proteins stereochemically optimized in size, stability, and folding speed. Biopolymers 83(2):122–134. https://doi.org/10.1002/bip.20537

Katragadda M, Magotti P, Sfyroera G, Lambris JD (2006) Hydrophobic effect and hydrogen bonds account for the improved activity of a complement inhibitor, compstatin. J Med Chem 49(15):4616–4622. https://doi.org/10.1021/jm0603419

Katz MD, Erstad BL (1989) Octreotide, a new somatostatin analogue. Clin Pharm 8(4):255–273

Kauer JC, Erickson-Viitanen S, Wolfe HR Jr, DeGrado WF (1986) p-Benzoyl-L-phenylalanine a new photoreactive amino acid photolabeling of calmodulin with a synthetic calmodulin-binding peptide. J Biol Chem. 261(23):10695–10700

Kawai T, Sun B, Yoshino H, Feng D, Suzuki Y, Fukazawa M, Nagao S, Wainscott DB, Showalter AD, Droz BA, Kobilka TS, Coghlan MP, Willard FS, Kawabe Y, Kobilka BK, Sloop KW (2020) Structural basis for GLP-1 receptor activation by LY3502970, an orally active nonpeptide agonist. Proc Natl Acad Sci U S A 117(47):29959–29967. https://doi.org/10.1073/pnas.2014879117

Knudsen LB, Nielsen PF, Huusfeldt PO, Johansen NL, Madsen K, Pedersen FZ, Thogersen H, Wilken M, Agerso H (2000) Potent derivatives of glucagon-like peptide-1 with pharmacokinetic properties suitable for once daily administration. J Med Chem 43(9):1664–1669. https://doi.org/10.1021/jm9909645

Kowalczyk W, Sobolewski D, Prahl A, Derdowska I, Borovickova L, Slaninova J, Lammek B (2007) The effects of N-terminal part modification of arginine vasopressin analogues with 2-aminoindane-2-carboxylic acid: a highly potent V2 agonist. J Med Chem 50(12):2926–2929. https://doi.org/10.1021/jm070174s

Kumar V, Lee JD, Clark RJ, Noakes PG, Taylor SM, Woodruff TM (2020) Preclinical pharmacokinetics of complement C5a receptor antagonists PMX53 and PMX205 in mice. ACS Omega 5(5):2345–2354. https://doi.org/10.1021/acsomega.9b03735

Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, Zhang Q, Shi X, Wang Q, Zhang L, Wang X (2020) Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature 581(7807):215–220. https://doi.org/10.1038/s41586-020-2180-5

Lau JL, Dunn MK (2018) Therapeutic peptides: historical perspectives, current development trends, and future directions. Bioorg Med Chem 26(10):2700–2707. https://doi.org/10.1016/j.bmc.2017.06.052

Li SC, Goto NK, Williams KA, Deber CM (1996) Alpha-helical, but not beta-sheet, propensity of proline is determined by peptide environment. Proc Natl Acad Sci U S A 93(13):6676–6681. https://doi.org/10.1073/pnas.93.13.6676

Li X, Yang D (2006) Peptides of aminoxy acids as foldamers. Chem Commun (camb) 32:3367–3379. https://doi.org/10.1039/b602230h

Lien S, Lowman HB (2003) Therapeutic peptides. Trends Biotechnol 21(12):556–562. https://doi.org/10.1016/j.tibtech.2003.10.005

Liu H, Kim HR, Deepak R, Wang L, Chung KY, Fan H, Wei Z, Zhang C (2018) Orthosteric and allosteric action of the C5a receptor antagonists. Nat Struct Mol Biol 25(6):472–481. https://doi.org/10.1038/s41594-018-0067-z

Lu Y, Freeland S (2006) On the evolution of the standard amino-acid alphabet. Genome Biol 7(1):102. https://doi.org/10.1186/gb-2006-7-1-102

Magana M, Pushpanathan M, Santos AL, Leanse L, Fernandez M, Ioannidis A, Giulianotti MA, Apidianakis Y, Bradfute S, Ferguson AL, Cherkasov A, Seleem MN, Pinilla C, de la Fuente-Nunez C, Lazaridis T, Dai T, Houghten RA, Hancock REW, Tegos GP (2020) The value of antimicrobial peptides in the age of resistance. Lancet Infect Dis 20(9):e216–e230. https://doi.org/10.1016/S1473-3099(20)30327-3

Mahalakshmi R, Balaram P (2006) The use of D-amino acids in peptide design. D-Amino Acids: A New Frontier in Amino Acid and Protein Research—Practical Methods and Protocols:415–430

Mathur P, Jagannathan NR, Chauhan VS (2007) Alpha, beta-dehydrophenylalanine containing cecropin-melittin hybrid peptides: conformation and activity. J Pept Sci 13(4):253–262. https://doi.org/10.1002/psc.841

Merutka G, Dyson HJ, Wright PE (1995) 'Random coil' 1 H chemical shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG. J Biomol NMR 5(1):14–24

Miller WR, Bayer AS, Arias CA (2016) Mechanism of action and resistance to daptomycin in *Staphylococcus aureus* and *Enterococci*. Cold Spring Harb Perspect Med 6(11):a026997. https://doi.org/10.1101/cshperspect.a026997

Mishra R, Behera LM, Rana S (2022) Binding of raloxifene to human complement fragment 5a ((h)C5a): a perspective on cytokine storm and COVID19. J Biomol Struct Dyn 40(3):982–994. https://doi.org/10.1080/07391102.2020.1820381

Mishra R, Rana S (2019) A rational search for discovering potential neutraligands of human complement fragment 5a ((h)C5a). Bioorg Med Chem 27(19):115052. https://doi.org/10.1016/j.bmc.2019.115052

Morrisett JD, David JS, Pownall HJ, Gotto AM Jr (1973) Interaction of an apolipoprotein (apoLP-alanine) with phosphatidylcholine. Biochemistry 12(7):1290–1299. https://doi.org/10.1021/bi00731a008

Peng SB, Zhang X, Paul D, Kays LM, Gough W, Stewart J, Uhlik MT, Chen Q, Hui YH, Zamek-Gliszczynski MJ, Wijsman JA, Credille KM, Yan LZ (2015) Identification of LY2510924, a novel cyclic peptide CXCR4 antagonist that exhibits antitumor activities in solid tumor and breast cancer metastatic models. Mol Cancer Ther 14(2):480–490. https://doi.org/10.1158/1535-7163.MCT-14-0850

Pollegioni L, Servi S (2012) Unnatural amino acids. Methods Mol Biol

Pomplun S (2020) Targeting the SARS-CoV-2-spike protein: from antibodies to miniproteins and peptides. RSC Med Chem 12(2):197–202. https://doi.org/10.1039/d0md00385a

Purcell AW, McCluskey J, Rossjohn J (2007) More than one reason to rethink the use of peptides in vaccine design. Nat Rev Drug Discov 6(5):404–414. https://doi.org/10.1038/nrd2224

Rader AFB, Reichart F, Weinmuller M, Kessler H (2018) Improving oral bioavailability of cyclic peptides by *N*-methylation. Bioorg Med Chem 26(10):2766–2773. https://doi.org/10.1016/j.bmc.2017.08.031

Ramesh VV, Roy A, Vijayadas KN, Kendhale AM, Prabhakaran P, Gonnade R, Puranik VG, Sanjayan GJ (2011) Conformationally rigid aromatic amino acids as potential building blocks for abiotic foldamers. Org Biomol Chem 9(2):367–369. https://doi.org/10.1039/c0ob00593b

Rana S, Kundu B, Durani S (2005) A small peptide stereochemically customized as a globular fold with a molecular cleft. Chem Commun (camb) 2:207–209. https://doi.org/10.1039/b413802c

Rana S, Sahoo AR (2015) Model structures of inactive and peptide agonist bound C5aR: insights into agonist binding, selectivity and activation. Biochem Biophys Rep 1:85–96. https://doi.org/10.1016/j.bbrep.2015.03.002

Rana S, Sahoo AR, Majhi BK (2016) Structural complexes of the agonist, inverse agonist and antagonist bound C5a receptor: insights into pharmacology and signaling. Mol Biosyst 12(5):1586–1599. https://doi.org/10.1039/c6mb00031b

Ree R, Varland S, Arnesen T (2018) Spotlight on protein *N*-terminal acetylation. Exp Mol Med 50(7):1–13. https://doi.org/10.1038/s12276-018-0116-z

Renfrew PD, Choi EJ, Bonneau R, Kuhlman B (2012) Incorporation of noncanonical amino acids into rosetta and use in computational protein-peptide interface design. PLoS ONE 7(3):e32637. https://doi.org/10.1371/journal.pone.0032637

Revilla-Lopez G, Torras J, Curco D, Casanovas J, Calaza MI, Zanuy D, Jimenez AI, Cativiela C, Nussinov R, Grodzinski P, Aleman C (2010) NCAD, a database integrating the intrinsic conformational preferences of non-coded amino acids. J Phys Chem B 114(21):7413–7422. https://doi.org/10.1021/jp102092m

Sachdeva S (2016) Peptides as 'Drugs': the Journey so far. Int J Pept Res Ther 23(1):49–60. https://doi.org/10.1007/s10989-016-9534-8

Saghyan AS, Langer P (2016) Non-proteinogenic α-Amino acids, natural origin, biological functions. In: Asymmetric synthesis of non-proteinogenic amino acids, pp 1–23. https://doi.org/10.1002/9783527804498.ch1

Sahoo AR, Mishra R, Rana S (2018) The Model Structures of the Complement Component 5a Receptor (C5aR) Bound to the Native and Engineered (h)C5a. Sci Rep 8(1):2955. https://doi.org/10.1038/s41598-018-21290-4

Schiller PW, Weltrowska G, Dung NTM, Lemieux C, Chung NN, Marsden BJ, Wilkes BC (1991) Conformational restriction of the phenylalanine residue in a cyclic opioid peptide analog: effects on receptor selectivity and stereospecificity. J Med Chem 34(10):3125–3132

Sebokova E, Christ AD, Wang H, Sewing S, Dong JZ, Taylor J, Cawthorne MA, Culler MD (2010) Taspoglutide, an analog of human glucagon-like Peptide-1 with enhanced stability and in vivo potency. Endocrinology 151(6):2474–2482. https://doi.org/10.1210/en.2009-1459

Sharma SP, Sharma J, Kanwar SS, Chauhan VS (2012) In vitro antibacterial and antimalarial activity of dehydrophenylalanine-containing undecapeptides alone and in combination with drugs. Int J Antimicrob Agents 39(2):146–152. https://doi.org/10.1016/j.ijantimicag.2011.10.008

Soundararajan M, Willard FS, Kimple AJ, Turnbull AP, Ball LJ, Schoch GA, Gileadi C, Fedorov OY, Dowler EF, Higman VA, Hutsell SQ, Sundstrom M, Doyle DA, Siderovski DP (2008) Structural diversity in the RGS domain and its interaction with heterotrimeric G protein alpha-subunits. Proc Natl Acad Sci U S A 105(17):6457–6462. https://doi.org/10.1073/pnas.0801508105

Squires KE, Montanez-Miranda C, Pandya RR, Torres MP, Hepler JR (2018) Genetic analysis of rare human variants of regulators of G protein signaling proteins and their role in human physiology and disease. Pharmacol Rev 70(3):446–474. https://doi.org/10.1124/pr.117.015354

Stevenazzi A, Marchini M, Sandrone G, Vergani B, Lattanzio M (2014) Amino acidic scaffolds bearing unnatural side chains: an old idea generates new and versatile tools for the life sciences. Bioorg Med Chem Lett 24(23):5349–5356. https://doi.org/10.1016/j.bmcl.2014.10.016

Stone TA (2017) Deber CM Therapeutic design of peptide modulators of protein-protein interactions in membranes. Biochim Biophys Acta Biomembr 1859(4):577–585. https://doi.org/10.1016/j.bbamem.2016.08.013

Tran TT, Treutlein H, Burgess AW (2006) Designing amino acid residues with single-conformations. Protein Eng Des Sel 19(9):401–408. https://doi.org/10.1093/protein/gzl024

Upert G, Luther A, Obrecht D, Ermert P (2021) Emerging peptide antibiotics with therapeutic potential. Med Drug Discov 9:100078. https://doi.org/10.1016/j.medidd.2020.100078

Vaiana AC, Neuweiler H, Schulz A, Wolfrum J, Sauer M, Smith JC (2003) Fluorescence quenching of dyes by tryptophan: interactions at atomic detail from combination of experiment and computer simulation. J Am Chem Soc 125(47):14564–14572. https://doi.org/10.1021/ja036082j

Vasudev PG, Chatterjee S, Shamala N, Balaram P (2011) Structural chemistry of peptides containing backbone expanded amino acid residues: conformational features of beta, gamma, and hybrid peptides. Chem Rev 111(2):657–687. https://doi.org/10.1021/cr100100x

Vlieghe P, Lisowski V, Martinez J, Khrestchatisky M (2010) Synthetic therapeutic peptides: science and market. Drug Discov Today 15(1–2):40–56. https://doi.org/10.1016/j.drudis.2009.10.009

Wishart DS, Sykes BD, Richards FM (1992) The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. Biochemistry 31(6):1647–1651

Zhang X, Boyar W, Toth MJ, Wennogle L, Gonnella NC (1997) Structural definition of the C5a C terminus by two-dimensional nuclear magnetic resonance spectroscopy. Proteins 28(2):261–267. https://doi.org/10.1002/(sici)1097-0134(199706)28:2%3c261::aid-prot13%3e3.0.co;2-g