A new amino acid for improving permeability and solubility in macrocyclic peptides through side chain-to-backbone hydrogen bonding

Jaru Taechalertpaisarn[†], Satoshi Ono[‡], Okimasa Okada[‡], Timothy C. Johnstone[†], R. Scott Lokey*[,†]

[†]Department of Chemistry and Biochemistry, University of California, Santa Cruz, 1156 High Street, Santa Cruz, California 95064, United States

[‡]Modality Laboratories, Innovative Research Division, Mitsubishi Tanabe Pharma Corporation, 1000 Kamoshida-cho, Aoba-ku, Yokohama, Kanagawa 227-0033, Japan

ABSTRACT: Despite the notoriously poor membrane permeability of peptides in general, many cyclic peptide natural products show high passive membrane permeability and potently inhibit a variety of "undruggable" intracellular targets. A major impediment to designing cyclic peptides with good permeability is the high desolvation energy associated with the peptide backbone amide NH groups. While several strategies have been proposed to mitigate this deleterious effect, only few studies have used polar side chains to sequester backbone NH groups. We investigated the ability of *N*,*N*-pyrrolidinyl glutamine (Pye), whose side chain contains a powerful hydrogen bond accepting C=O amide group but no hydrogen bond donors, to sequester exposed backbone NH groups in a series of cyclic hexapeptide diastereomers. Analyses revealed that specific Leu-to-Pye substitutions conferred dramatic improvements in aqueous solubility and permeability in a scaffold- and position-dependent manner. Therefore, this approach offers a complementary tool for improving membrane permeability and solubility in cyclic peptides.

Introduction

Advances in medicine and molecular cell biology continue to generate therapeutic targets whose large, flat binding interfaces make them challenging to drug with traditional, Rule of 5 (Ro5)-compliant small molecules.[1-6] Although large biomolecules can often bind with high affinity to such undruggable sites, due to their inability to cross the cell membrane these drugs are generally restricted to parenteral delivery against extracellular targets.[7-9] Thus the inverse relationship between molecular weight and cell permeability has focused the search for new chemical

matter against undruggable intracellular targets within the chemical space that lies between small molecules and biologics, where there are a growing number of drugs, leads, and model systems that can engage classically undruggable sites yet exhibit the favorable membrane permeability and oral absorption typical of small molecule drugs.[10-13] Among the chemotypes that are in the “beyond Rule of 5 (bRo5)” chemical space, cyclic peptides are gaining attention as they can potentially interact with protein targets but still present some of the physicochemical properties of traditional small molecule drugs. The influence of emerging bRo5 scaffolds has already driven novel approved oral drugs toward larger molecular weights over the past two decades.[14, 15]

Cyclic peptides have several favorable pharmacological characteristics as therapeutic agents and chemical probes.[16-18] Cyclization of peptides reduces conformational flexibility, enhances proteolytic stability and increases binding affinity and selectivity.[16-18] The relatively larger surface areas of cyclic peptides, compared to conventional small molecules, allow them to engage large protein surfaces, leading to a high affinity toward targets that lack well-defined binding pockets.[18-20] Moreover, cyclic peptides can be diversified enormously by the addition of non-proteinogenic amino acids (e.g., D-amino acids, *N*-methyl amino acids, peptoids, etc.) and by varying ring size and backbone linkages (e.g., lariats, depsipeptides, inclusion of polyketide fragments).

Lessons learned from the orally bioavailable cyclic peptide cyclosporine A (CsA)[9, 21-23], along with numerous other model systems[17, 19, 24], have confirmed the importance of overall lipophilic character, as well as the degree to which hydrogen bond donors can be shielded from solvent, in determining membrane permeability. Several methods have been proposed to minimize exposed amide protons, such as (i) stabilization of intramolecular hydrogen bonding through stereochemistry of amino acid residues[24-28], (ii) replacement of polar NH residues that cannot participate in intramolecular hydrogen bonding (IMHB) with *N*-methyl amino acids or peptoids[9, 29-33], (iii) α-methylation of amino acid residues to reinforce the IMHB network[34] and (iv) steric shielding with nearby hydrophobic side chains.[1, 30, 35, 36]

While these methods (which have been summarized in a recent review[18]) successfully enhance membrane permeability, some drawbacks impede their application. *N*-methylation removes amide protons that can potentially interact with protein targets, and *N*-methylation at some residues may disrupt IMHB networks leading to a decrease in membrane permeability.[30] Furthermore, steric occlusion of HBD requires the introduction of bulky side chains, which can impede target binding and increase overall lipophilicity beyond acceptable limits. Consequently, new strategies for sequestering polar amide protons that are orthogonal to N-methylation and steric occlusion would be useful.

Our interest in the use of side chain hydrogen bond acceptors (HBAs) to sequester exposed amide hydrogen bond donors (HBDs) was motivated in part by the observation that these types of side chain-to-backbone (SC-BB) interactions are present in some crystal structures of natural product cyclic peptides that are known to have intracellular targets and whose structures suggest that they may be passively permeable. For example, argyrin B, [37] which inhibits the prokaryotic and mitochondrial elongation factor G (EF-G), contains an IMHB between the MeO- of its 4'-methoxytryptophan and its backbone amide NH (Figure 1A).[38]

A few synthetic cyclic peptides have been reported that utilize exocyclic HBAs to sequester exposed backbone amides. Thansandote, *et al*.,[34] incorporated a 2-pyridinylalanine side chain into a cyclic hexapeptide, which improved solubility but failed to improve permeability, possibly due to the deleterious effect of the pyridyl side chain on hydrocarbon-water partitioning. Another system reported by Yudin, *et al*., integrated an exocyclic amide motif into a peptide macrocycle, which formed hydrogen bonds with two backbone NH groups, resulting in an improvement of passive membrane permeability. However, the exocyclic amide introduced an additional hydrogen bond, thus potentially limiting the generality of this approach.[39-42] Although other cyclic peptide systems have been described that show side chain HBAs interacting with backbone amide NH groups, the impact of these interactions on the compounds' drug-like properties were not investigated systematically.[43-45]

Herein, we set out to investigate the interaction between side chain HBAs and backbone amides in a series of cyclic hexapeptide model systems, and, in turn, evaluate their effect on conformation and their ability to improve membrane permeability. By substituting a side chain HBA on each position in three different congeneric series, we found that side chain HBAs can sequester exposed amide protons by forming SC-BB H-bonds, supported by X-ray crystal and NMR solution structures. As a result, we observed an increase in lipophilicity and membrane permeability compared to isomeric scaffolds in which the position and stereochemistry of the polar side chain did not favor SC-BB H-bonding. Moreover, the substitution of a polar, HBA-containing side chain dramatically increased aqueous solubility, thereby offsetting the combined deleterious effect of multiple lipophilic residues.

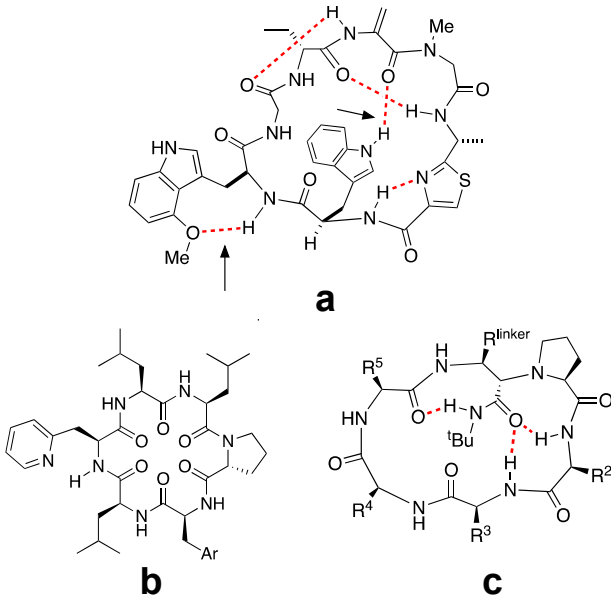


Figure 1. Structures of (**a**) argyrin B, (**b**) Simpson's 2-pyridinylalanine cyclic hexapeptide and (**c**) Yudin's 18-membered ring macrocycle with exocyclic amide. Hydrogen bonds are indicated with dashed red lines.

Results and Discussion

Side chain-to-backbone Hydrogen Bonding of Model Structures

In order to identify side chains that could sequester backbone amide NH groups through IMHB interactions, we performed computational studies on a series of model compounds bearing various side-chain HBAs. We performed conformational searches in an implicit, low-dielectric solvent continuum to mimic the membrane dielectric, and plotted the distance between the side chain HBA and backbone amide NH vs. energy, relative to the lowest-energy conformation (Figure S1). The results showed that *N*,*N*-disubstituted Gln residues had low-energy conformers with short HBA-HBD distances (< 2.0 Å; Figure S1a, d, e), while the distances from *N*,*N*-disubstituted Asn were slightly longer (Figure S1f, h). Similar HBA-HBD distances were observed when replacing an amide functional group with an ester (Figure S1c).

The encouraging *in silico* data prompted us to investigate the degree of SC-BB H-bonding in the set of synthetic model structures **4** shown in Table 1. For commercially available amino acids (**4-Ala**, **4-Val**, **4-Leu**, **4-Ile**, **4-Gln**, **4-Glu(OMe)**), Boc-protected amino acids were esterified with 4-phenyl-1-butanol, deprotected, and acetylated to obtain compounds **4** (Scheme S1a). Substituted amides (**4-Gln(NMe$_2$)**, **4-Gln(Pyr)**, **4-Asn(Pyr)**, **4-Asn(NMe$_2$)**) were synthesized from corresponding *N*- and *C*-protected Glu or Asp amino acids (Boc-Glu-O^{t}Bu and Boc-Asp-

O^{t}Bu; Scheme S1b). After amide coupling with pyrrolidine (Pyr) or *N*,*N*-dimethylamine, the intermediates were treated in TFA to deprotect all acid-labile protecting groups, the Boc-protecting group was reintroduced, and the compounds were carried forward as described above.

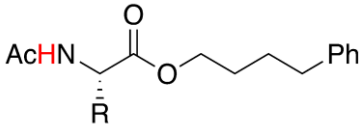
NMR chemical shifts (δ) and temperature coefficients ($\Delta\delta_{NH}/\Delta T$) are widely used to gain structural information in peptides and proteins.[26, 46, 47] Temperature coefficients of backbone amide protons have been used to report on their degree of solvent exposure; for well-structured proteins in aqueous solution, amide NH protons with temperature shifts above (more positive than) -4.0 ppb/K are considered to be involved in IMHB, whereas more negative values indicate greater solvent exposure.[46] This relationship has been attributed to the strong ($1/r^3$) dependence of NH chemical shifts on the NH–O distance, where the NH becomes more deshielded and shifts downfield as the H-bond becomes shorter.[48]

In proteins, solvent-exposed amides have relatively large thermal shifts due to the weaker H-bond to water compared to the NH–to–carbonyl H-bond. However, in nonpolar solvents[49, 50], fully exposed NH groups can have small temperature shifts because they interact with neither solvent nor acceptor atoms within the molecule. For NH groups that are involved in IMHB in nonpolar solvents, temperature shifts can either be large, if the IMHB interaction has a strong negative entropic component, e.g., with multiple rotatable bonds between the donor and acceptor, or small, if the enthalpic gain outweighs the entropic penalty of H-bond formation and the interaction remains stable with increasing temperature. Therefore, interpreting thermal shift data is challenging, especially in small conformationally dynamic peptides, where additional information such as NH chemical shifts derived from model systems can be useful.[51]

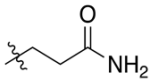
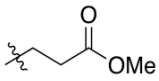
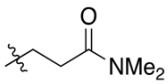
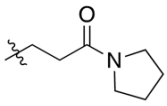
To avoid artifacts due to *intermolecular* H-bonding, we determined the NH chemical shifts of **4-Ala** and **4-Gln(Pyr)** as a function of concentration (Figure S4). At concentrations below 50 mM the NH chemical shift remained nearly constant, indicating minimal intermolecular H-bonding. The model compounds with aliphatic side chains (**4-Ala**, **4-Val**, **4-Leu**, **4-Ile**; Table 1 and Figure S3) showed chemical shifts of ~6.0 ppm and temperature shift coefficients between -2.95 and -3.54 ppb/K. These values are consistent with observations from Gellman, *et al.,* who saw comparable chemical shifts and temperature coefficients in CD_2Cl_2 for similar model systems lacking in IMHB.[50] For the substituted Gln amides (**4-Gln(NMe$_2$)**, **4-Gln(Pyr)**), the amide proton chemical shifts appeared approximately 1 ppm downfield of the NH protons in the aliphatic controls and showed significantly larger negative $\Delta\delta_{NH}/\Delta T$ values (-6.48 and -7.84 ppb/K). Both the chemical shifts and temperature coefficients of ester **4-Glu(OMe)** were similar to the aliphatic controls, consistent with the weaker hydrogen bond-accepting ability of esters

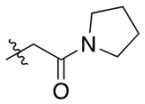
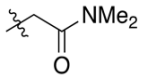
compared to amides.[50] Taken together, these results are consistent with previous studies finding comparably upfield chemical shifts in nonpolar media for weakly hydrogen-bonding NH groups, and downfield chemical shifts for NH groups that are significantly hydrogen-bonded. Interestingly, while the chemical shifts of substituted Asn amides (**4-Asn(Pyr)**, **4-Asn(NMe$_2$)**) were similarly downfield (6.81 and 6.89 ppm) to the Gln compounds, their $\Delta\delta_{NH}/\Delta T$ values (-1.66 and -1.52 ppb/K) were much less negative than either the Gln derivatives or the aliphatic controls. The larger temperature coefficients of the Gln compared to Asn residues are consistent with the larger entropic penalty for H-bond formation of the more flexible Gln side chain. Combined with the computational studies, these NMR data prompted us to investigate the **4-Gln(Pyr)** residue, which we abbreviate as "Pye" (pyrrolidine amide of glutamic acid 'E'), as the most promising candidate for eliciting permeability-enhancing SC-BB interactions in cyclic peptides.

Table 1. Chemical Shifts and Amide Temperature Coefficients of Model Structures in $CDCl_3$[a].



4

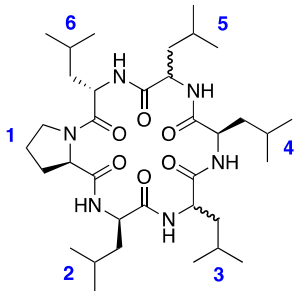
ID	R	δ_{NH} (300 K)	$\Delta\delta_{NH}/\Delta T$ (ppb/K)
4-Ala	Me	6.08	-3.54
4-Val		5.97	-3.31
4-Leu		5.86	-3.02
4-Ile		5.96	-2.95
4-Gln		6.59	-5.37
4-Glu(OMe)		6.18	-3.51
4-Gln(NMe$_2$)		6.88	-6.48
4-Gln(Pyr) (Pye)		7.11	-7.84
4-Asn(Pyr)		6.81	-1.66

			
4-Asn(NMe_2)		6.89	-1.52

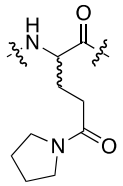
[a]All experiments were performed at a concentration of 50 mM or below.

NMR Studies of Cyclic Hexapeptides

Both L and D-Pye were synthesized on a multigram scale (Scheme S2) first using EDC and HOBT to couple pyrrolidine to the side chain of Fmoc-Glu-O^{t}Bu, followed by removal of the *tert*-butyl ester with TFA and diethyl ether precipitation. Previously we identified a series of diastereomeric cyclic hexapeptide scaffolds whose membrane permeabilities[52] varied significantly depending on their ability to sequester exposed amide NH groups in IMHB.[53] In order to determine the impact of Pye on passive permeability, we selected three of the least permeable diastereomers, **1**, **2** and **3**, which differ in stereochemistry at positions Leu3 and Leu5 (we replaced the Tyr6 residue from the previous study with Leu to avoid the deleterious effect of the phenolic OH on permeability). Herein, compounds are named by the type of polar residue (Asn, Pye, etc) preceded by the scaffold number and followed by a superscript indicating the position of the substitution (e.g., **1-Pye^2** refers to scaffold **1** in which Leu2 is replaced with Pye). Substituting each Leu with Pye in the three scaffolds yielded a total of 18 compounds, including the three parental compounds **1**, **2** and **3**.



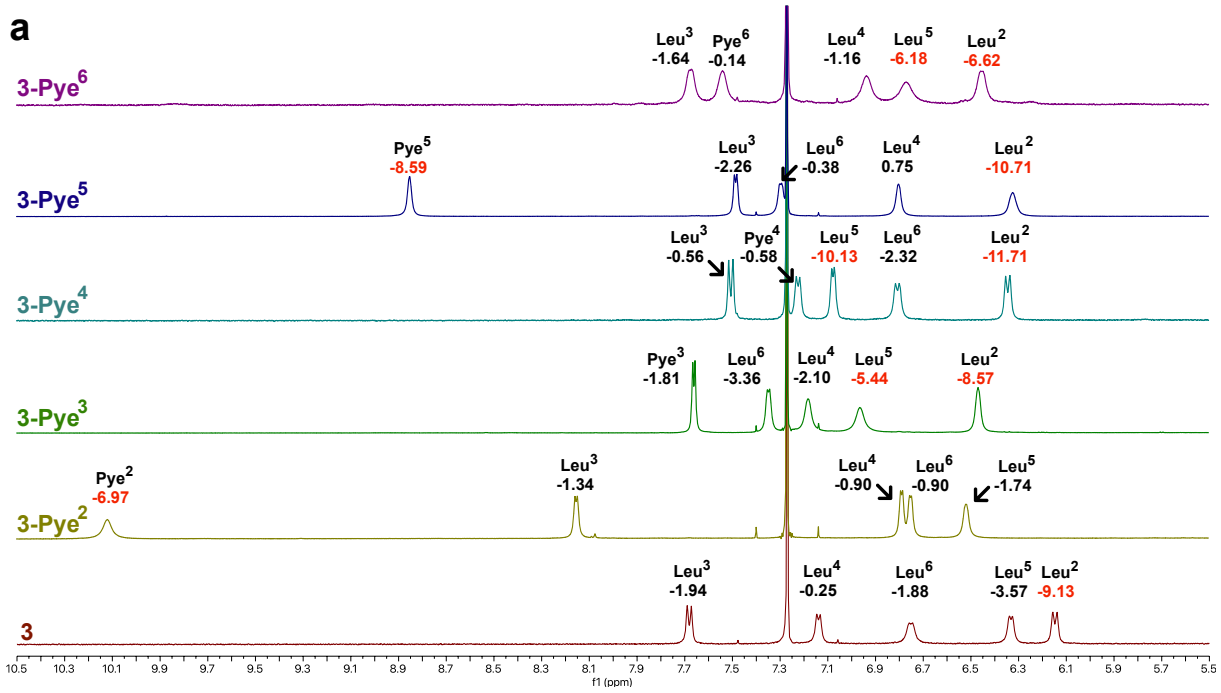
ID	Residue chirality					
	1	2	3	4	5	6
1	D	D	D	D	L	L
2	D	D	D	D	D	L
3	D	D	L	D	L	L

Pye = 

For the series based on **1** and **2**, we were unable to assign all of the resonances due to line broadening and/or peak overlap. However, for the series based on diastereomer **3**, the peaks were sharp and dispersed enough to allow complete assignment of the parental compound as well as all of the Pye positional variants (Figure 2). For **3-Pye^2** and **3-Pye^5** (Figure 2), the amide proton of the Pye residue fell significantly downfield of the other NH resonances, even as far as 10 ppm for **3-Pye^2**. Based on the chemical shifts and temperature coefficients of the

model peptides (Table 1), we hypothesized that the Pye side chain in these compounds form IMHB to their own backbone amide NH groups. For many of the other compounds, including **3-Pye^3**, **3-Pye^4**, and **3-Pye^6** (Figure 2), (as well as **1-Pye^4**, **1-Pye^5**, and **1-Pye^6**) (Figure S5b), the chemical shift of the Pye NH was in the same range as that of the other amide NH resonances, between 6 and 8.5 ppm.

The temperature coefficients of **1**, **2** and **3**, showed at least one amide proton exposed to solvent (Figure 2 and S4b, d); for the fully assigned compound **3**, the NH with the high temperature coefficient corresponded to Leu2 (-9.13 ppb/K). For all other **3-Pye** positional variants (excluding **3-Pye^2**, in which Leu2 is replaced with Pye), the Leu2 NH was the most shielded and also had the largest temperature coefficient compared to the other NH peaks. For **3-Pye^3**, **3-Pye^4**, and **3-Pye^6**, Leu5 also showed a relatively high temperature coefficient. The high variability in both chemical shifts and temperature coefficients indicates that the position of the Pye residue in the scaffold has a significant effect on its ability to participate in IMHB and demonstrates that substitution of a single aliphatic residue with an HBA-containing residue can have a strong position-dependent effect on the scaffold's low-dielectric conformation.



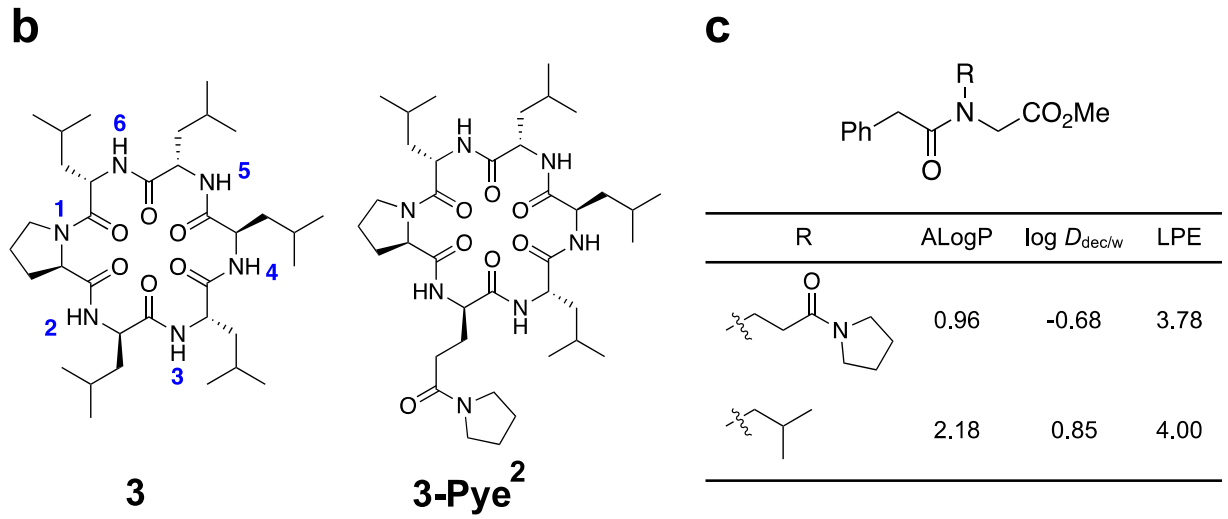


Figure 2. (**a**) ^{1}H NMR of amide protons of **3** series at a concentration of ~10 mM. Values under labels indicate the $\Delta\delta_{NH}/\Delta T$ in ppb/K. Amide temperature coefficients less than -4.0 are highlighted in red. (**b**) Structure of **3** and **3-Pye^{2}**. (**c**) The intrinsic effect of Pye and isobutyl side chains on LPE in the absence of HBD.

Lipophilicity, Passive Permeability and Solubility of Pye-scanning Cyclic Hexapeptides

Previously we introduced a metric called lipophilic permeability efficiency (LPE) for assessing the efficiency with which a compound achieves membrane permeability at a given lipophilicity.[53] LPE is defined as log $D_{dec/w}$ - 1.06(ALogP) + 5.47, where log $D_{dec/w}$ is the experimental partition coefficient between 1,9-decadiene and PBS buffer at pH 7.4, and ALogP[54] is a calculated, atomistic 2D octanol/water partition coefficient (a slight variation on the more commonly used cLogP). Hydrocarbons provide a good model of the membrane interior, and thus partitioning between water and 1,9-decadiene reflects the energic penalty of desolvating exposed polar functionality, in particular HBDs, as they cross the barrier region of the phospholipid bilayer. Therefore, for compounds that are water-soluble, log $D_{dec/w}$ correlates positively with passive permeability. On the other hand, since octanol is capable of forming hydrogen bonds with solute, ALogP does not penalize exposed HBD but rather reflects a compound's minimum lipophilic character in the aqueous environment. Therefore, for highly lipophilic compounds (with ALogP above ~4), ALogP shows a good negative correlation with solubility as well as a negative correlation with permeability in the insoluble regime.[55-57] By subtracting ALogP from log $D_{dec/w}$, LPE thus provides a measure of a scaffold's ability to sequester polar functionality (in particular HBD) and therefore its potential to achieve high permeability at a given ALogP-defined lipophilicity.

To determine the ΔLPE of the Pye residue (relative to an aliphatic control) in the absence of IMHB, we synthesized a simple model compound in which the Pye functional group was introduced as a peptoid side chain (Figure 2c).

The Pye substitution decreased LPE by 0.22 units relative to the isobutyl control.[53] Next, LPE and thermodynamic aqueous solubilities were determined for parent compounds **1**-**3** and their Pye substitutions (Table 2). Replacing Leu with Pye decreases ALogP by 1.28 units. The same substitution also decreased log $D_{dec/w}$, by a variable amount depending on both the scaffold and the position of the Pye residue within the scaffold. For most compounds, substitution of Leu for Pye caused a significant decrease in LPE (by up to 1 LPE unit); however, for **1-Pye^2** and **3-Pye^2**, LPE *increased* upon Pye substitution. This observation is consistent with the upfield chemical shift and large temperature coefficient of Leu2 in parent compound **3**, indicating solvent exposure of the Leu2 NH. Although the NH resonances in parent compound **1** could not be assigned due to peak overlap, there is a single upfield amide peak with a large temperature coefficient that, based on the other similarities to scaffold **1**, likely corresponds to Leu2. The far downfield chemical shifts of the Pye residue in **1-Pye^2** and **3-Pye^2** suggest that the Pye side chain's C=O is able to sequester these exposed NH groups via an SC-BB IMHB. The temperature shift of the Pye2 amide in **1-Pye^2**, however, is small (-1.1 ppb/K), while that of **3-Pye^2** is much larger (-7.0 ppb/K). Since the Pye2 substitution's effect on LPE is greater for scaffold **1** than scaffold **3** (suggesting that the corresponding SC-BB IMHB is stronger for **1-Pye^2**), perhaps the SC-BB interaction in **1-Pye^2** is enthalpically favorable enough to offset the entropic penalty of H-bond formation, thus lowering the temperature coefficient for this amide.

Table 2. Pye-scanning Cyclic Hexapeptides Experimental Data

ID	log $D_{dec/w}$	LPE[a]	P_{app}[b]	solubility[c]
1	2.0	3.5	8.1	16
1-Pye^2	1.3	4.1	1.7	640
1-Pye^3	-0.41	2.5	0.66	550
1-Pye^4	0.02	2.9	0.30	580
1-Pye^5	-0.42	2.5	0.32	510
1-Pye^6	-0.30	2.6	0.66	720
2	2.0	3.5	5.3	240
2-Pye^2	-0.29	2.6	0.56	610
2-Pye^3	-0.29	2.6	1.1	450
2-Pye^4	0.040	2.9	1.3	460
2-Pye^5	-0.82	2.0	0.13	550
2-Pye^6	-0.29	2.6	1.5	560
3	2.0	3.6	8.0	63
3-Pye^2	1.1	3.9	1.5	730
3-Pye^3	-0.22	2.7	0.23	140
3-Pye^4	0.21	3.1	0.14	580
3-Pye^5	0.14	3.0	0.28	550
3-Pye^6	0.16	3.0	0.18	600

[a]LPE = log $D_{dec/w}$ – 1.06 x ALogP + 5.47[53]; ALogP = 3.74 for **1**, **2** and **3**; 2.46 for Pye-containing cyclic hexapeptides. [b]PAMPA x 10^{-6} cm/s, N=4. [c]μM, pH 7.4, N=3.

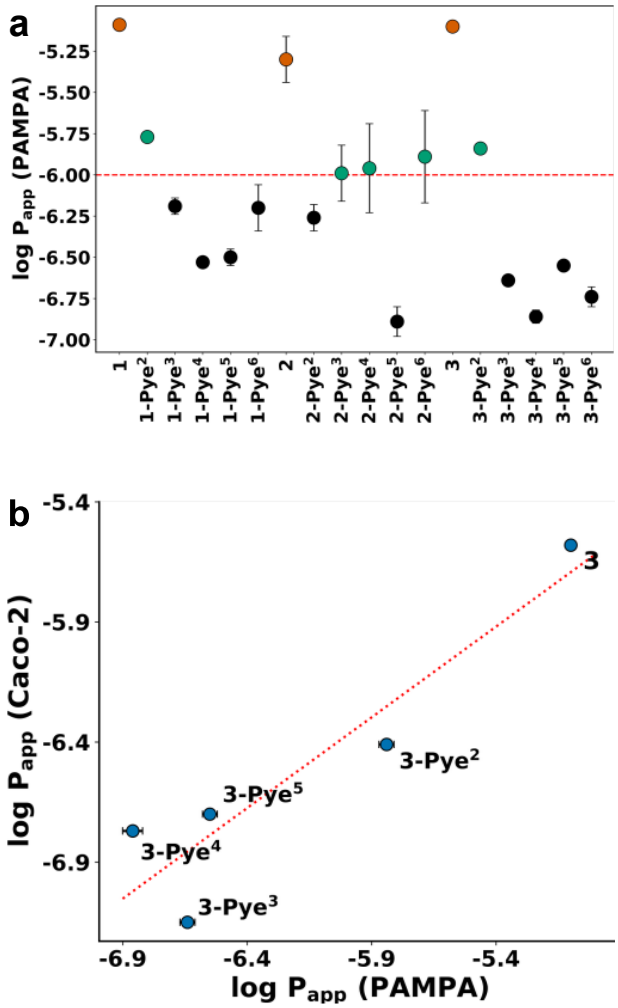


Figure 3. (**a**) PAMPA permeabilities of cyclic hexapeptides **1**, **2**, **3** and their derivatives. Minimum acceptable oral permeability is shown in dashed line**.** (**b**) PAMPA and Caco-2 experiments of compounds **3** and their derivatives show similar permeability rate (R^2 = 0.86). High and low permeability standards are measured along with test compounds in Caco-2 experiment to ensure the accuracy; N=4 for PAMPA experiment; N=1 for Caco-2 experiment.

The parallel artificial membrane permeability assay (PAMPA)[58] is a cell-free permeation tool for screening passive permeabilities (P_{app}). PAMPA measures the rate of diffusion of test compounds from donor to acceptor compartments across a solution of 1% lecithin in n-dodecane trapped in a polyvinyldifluoroethane membrane. The positive correlation between permeation and lipophilicity generally holds for compounds in the soluble regime, generally below ALogP ~ 4. At higher lipophilicities, effective permeability decreases due to factors such as membrane sequestration and poor aqueous solubility, a trend that has been observed in both cyclic peptides and small molecule drugs.[27, 59, 60] Thus, a plot between log P_{app} and ALogP produces a downward-facing parabola with a vertex near ALogP ~ 4.[25]

Compounds **1**, **2** and **3** showed excellent PAMPA permeabilities but had modest solubilities, consistent with their relatively high lipophilicity (ALogP = 3.74, Table 2). Most of the Pye-substituted compounds had P_{app} values below 1 x 10^{-6} cm/s except **1-Pye^2**, **2-Pye^3**, **2-Pye^4**, **2-Pye^6** and **3-Pye^2** (Figure 3a). Permeabilities and LPEs of **1-Pye^2** and **3-Pye^2** were clearly higher than the other Pye-substitutions. In addition, selected compounds in the **3** series were measured for their rate of flux across the human colon carcinoma cell (Caco-2) line, and their permeability rates were similar to those observed in PAMPA (Figure 3b).

Besides their ability to sequester backbone NH groups, another advantage of polar side chains is their potential for increasing a compound's aqueous solubility. Indeed, in all cases Pye substitution led to a significant improvement in solubility over the parental cyclic hexapeptides (Table 2). Interestingly, the degree to which introduction of the Pye residue improved permeability was also dependent on scaffold and position. For example, when Leu6 was replaced with Pye on scaffold **1**, solubility improved 45-fold, whereas replacing Leu3 for Pye on scaffold **2** improved solubility only two-fold. Interestingly, the degree to which Pye substitution improved aqueous solubility did not correlate with the effect on LPE. For example, the two compounds that showed an increase in LPE upon Pye substitution (**1-Pye^2** and **3-Pye^2**) also saw substantial improvements in solubility, consistent with a degree to which structural/conformational factors that contribute to solubility and permeability can be considered independently.

Membrane Permeability of 3-Pye^2 and 3-Pye^3 Libraries

To investigate the effect of Pye on lipophilicity and membrane permeability in the context of various amino acid combinations, we synthesized **3-Pye^2** and **3-Pye^3** libraries using the split-pool approach in which two residues were mutated to different hydrophobic side chains (Figure 4a). We chose these two scaffolds as they had the largest log $D_{dec/w}$ difference between Pye positional variants on the same scaffold, and the ^{1}H NMR spectra in the scaffold **3** series were interpretable. We selected side chain combinations that would cover a broad ALogP range, in this case from 0.5 to 4.7. The experimental hydrocarbon/water coefficients shown in Figure 4b agreed with Pye-scanning data in Table 2, showing that the log $D_{dec/w}$ values for the compounds based on the **3-Pye^2** positional variant were greater than those based on **3-Pye^3**. Matched pairs bearing the same combination of side chains from the **3-Pye^2** and **3-Pye^3** libraries had log $D_{dec/w}$ values that differed on average by one log unit, highlighting the importance of the Pye residue's position in determining its effect on physico-chemical properties.

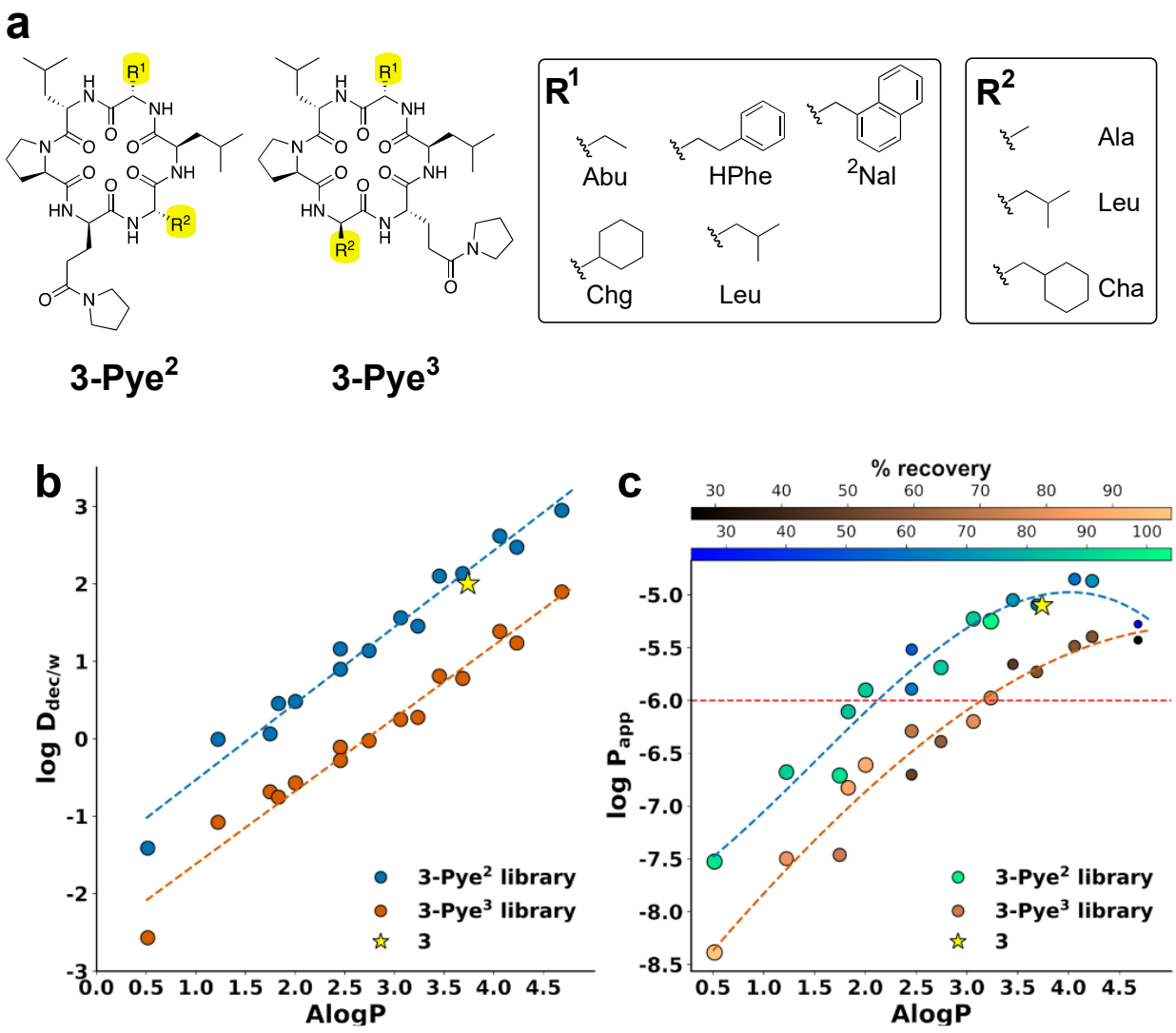


Figure 4. (**a**) Structures of **3-Pye²** and **3-Pye³** libraries and side chain composition. (**b**) Log $D_{dec/w}$ vs ALogP of **3-Pye²** and **3-Pye³** libraries. (**c**) Log P_{app} vs ALogP of **3-Pye²** and **3-Pye³** libraries. Red-dashed line indicates the minimum acceptable oral absorption. Compound **3** is shown as a yellow star.

The plot of PAMPA permeability (log P_{app}) versus ALogP for the **3-Pye²** and **3-Pye³** libraries revealed a similar trend, in which the **3-Pye²** derivatives were generally more permeable than the **3-Pye³** derivatives of the same ALogP (i.e., comparing pairs of compounds bearing the same side chain combinations; Figure 4c). At ALogP values below 4, the **3-Pye²** library was 0.5-1.0 log unit higher in permeability than the **3-Pye³** library. Above ALogP = 4, PAMPA permeabilities for both libraries showed the characteristic downward trend as aqueous solubilities fall below the threshold required to sustain permeability (Table S2-3). The two naphthyl-Ala derivatives **3-Pye²(Nal⁵)** and **3-Pye³(Nal⁵)** have nearly the same ALogP as the parent compound **3**, with the increased lipophilicity introduced by the Nal residue offsetting the decrease in lipophilicity upon substitution of Leu2 or Leu3 with Pye. But whereas the permeability of **3-Pye²(Nal⁵)** was comparable to that of **3**, the P_{app} of **3-Pye³(Nal⁵)** was 0.6 log units lower than the

P_{app} of **3**. In addition, two of the **3-Pye^2** derivatives (**Cha^{3}Chg^5** and **Cha^{3}HPhe^5**; ALogP = 4.1 and 4.2, respectively) achieved permeabilities greater than that of the all-Leu parent scaffold, whereas none of the **3-Pye^3**-based derivatives matched the permeability of the parent scaffold. Because of the lower LPE of the **3-Pye^3** scaffold, its derivatives require a higher ALogP to achieve the same of permeability as scaffolds of higher LPE (e.g.,**3-Pye^2**); in this case, the lipophilicity required for the **3-Pye^3** -based compounds to achieve the permeability of **3** puts them on the descending (low solubility) part of the P_{app} vs. ALogP curve (Figure 4c and Table S2 and S3).

Overall, the results from scaffold **3** suggest that substitution with Pye can enable SC-BB interactions which preserve the LPE of the parent scaffold, thus allowing for incorporation of larger, more lipophilic side chains without compromising water solubility. For example, **3-Pye^2(Cha^{3}Chg^5)** (ALogP = 4.1, P_{app} = 14 x 10^{-6} cm/s, solubility = 74 μM; Table S2) has better solubility and permeability than **3** (ALogP = 3.7, P_{app} = 8.0 x 10^{-6} cm/s, solubility = 63 μM; Table 2).

Structural Elucidation of 3-Pye^2 derivatives

The exceptional permeability of the **3-Pye^2** library encouraged us to explore the structural basis of the Pye residue's effect on permeability. We obtained structures by single crystal X-ray diffraction of two side chain variants on **3-Pye^2**, **3-Pye^2(Ala^{3}Nal^5)** and **3-Pye^2(HPhe^5)**, crystallized by vapor-diffusion in THF-pentane (Figure 5). These two crystals showed virtually identical backbone conformations (RMSD = 0.07 Å) and had nearly identical C_α-C_β vectors (RMSD = 0.09 Å). As expected from the NMR studies, the Pye2 side chain C=O showed a strong hydrogen bond to its own backbone amide NH (Figure 5a-b, black arrow). The distance and N–H--O angle were 1.9 Å and 155º for both crystals, respectively, which is an optimal hydrogen bond geometry.[50] Three other backbone amides participated in IMHB as β-turn motifs (i.e., 3rd NH to 6th CO and 6th NH to 3rd CO; Figure 5a-b). The NH groups of residues 4 and 5 formed *inter*molecular hydrogen bonds with other molecules in the crystal lattice (Figure S13).

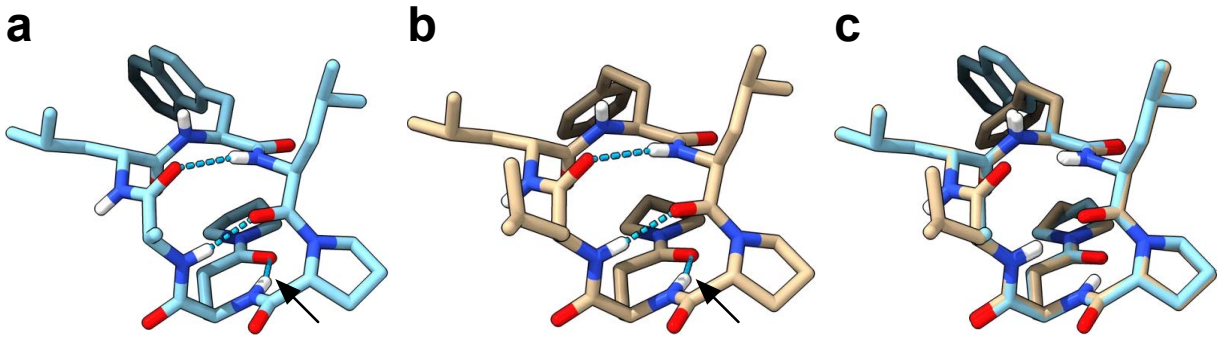


Figure 5. X-ray crystal structures of (**a**) **3-Pye^2(Ala^{3}Nal^5)** and (**b**) **3-Pye^2(HPhe^5)**. Hydrogen bond between Pye and 2nd amide is indicated with black arrow. Putative hydrogen bonds are indicated with dashed blue lines. (**c**) Overlay between **3-Pye^2(Ala^{3}Nal^5)** and **3-Pye^2(HPhe^5)**.

These crystal contacts suggested that these amide NH groups might be exposed in low-dielectric solvent; however, this hypothesis was inconsistent with the small $\Delta\delta_{NH}/\Delta T$ NMR temperature shifts of these NH groups (Figure S6) as well as the high LPE and PAMPA permeabilities of these compounds (Table S2). We hypothesized that crystal packing forces may impact their conformation, prompting us to investigate the solution structure of **3-Pye^2(Ala^{3}Nal^5)** in $CDCl_3$. By collecting NOESY spectra at various mixing times and performing a careful analysis of each crosspeak volume, we obtained a total of 20 crosspeaks that complied with the initial rate approximation.[61-63] The distance between the δ-proline geminal protons (1.78 Å) was used as a reference to calculate other interproton distances (Table S5). After applying 10% random noise to the calculated distances, 17 out of the 19 distances matched the crystal structure, and the two remaining distances were close to the predicted distance which suggested the overall structural similarity between the crystallographic and solution structures.

We did not detect a $C^{\alpha}H$-to-$C^{\alpha}H$ NOESY cross peak between Pro1 and Leu6, indicating that the Pro1 amide is most likely in the *trans* conformation. To obtain the solution NMR ensembles from the NOESY-derived distances, we performed an unrestrained conformational search using multicanonical molecular dynamics (McMD) simulations.[64] The 20 structures from the ensemble obtained by McMD at T = 300 K with the fewest violations from the NOESY-derived distances closely resembled the crystal structure (RMSD = 0.27 – 0.64 Å; Figure 6a-c) and were in a cluster that made up 9% of the unrestrained ensemble, although NOE violations averaged over the entire unrestrained McMD ensemble were also relatively low, indicating that the McMD simulation accurately captured the solution conformation in $CDCl_3$ (Table S5). The side chains were flexible except for the Pye2 side chain, which uniformly converged to form the expected SC-BB hydrogen bond.

The backbone NH groups of Ala3 and Leu6 formed the same transannular hydrogen bonding pattern as found in the crystal structure (Figure S14). The amides of Leu4 and Nal5, however, which formed crystal contacts in the X-ray structure, rotated inward in the solution ensemble to orientate toward the center of the macrocycle (Figure 6c). Specifically, the Leu4 NH reoriented to form a hydrogen bond with the C=O of Pye2, and the Nal5 NH reoriented to form a hydrogen bond with the C=O of Ala3 (Figure 6d-e), interactions that were absent in the crystal structure (Figure 6f). These reorientations around the 4th and 5th amide protons between the solid state and solution conformations amounted to a relatively subtle shift in overall backbone geometry, which is perhaps surprising given the net change in hydrogen bonding between the two structures. In solution all of the amide NH groups of **3-Pye^2(Ala^{3}Nal^5)** are involved in IMHB, consistent with the high LPE and PAMPA permeability of **3-Pye^2** and its derivatives. The large discrepancy in overall 3D polar surface area between the solid state and solution conformations despite similar backbone geometries suggests that dramatic shifts in 3D polar surface area can be modulated by relatively subtle shifts in amide rotamers.

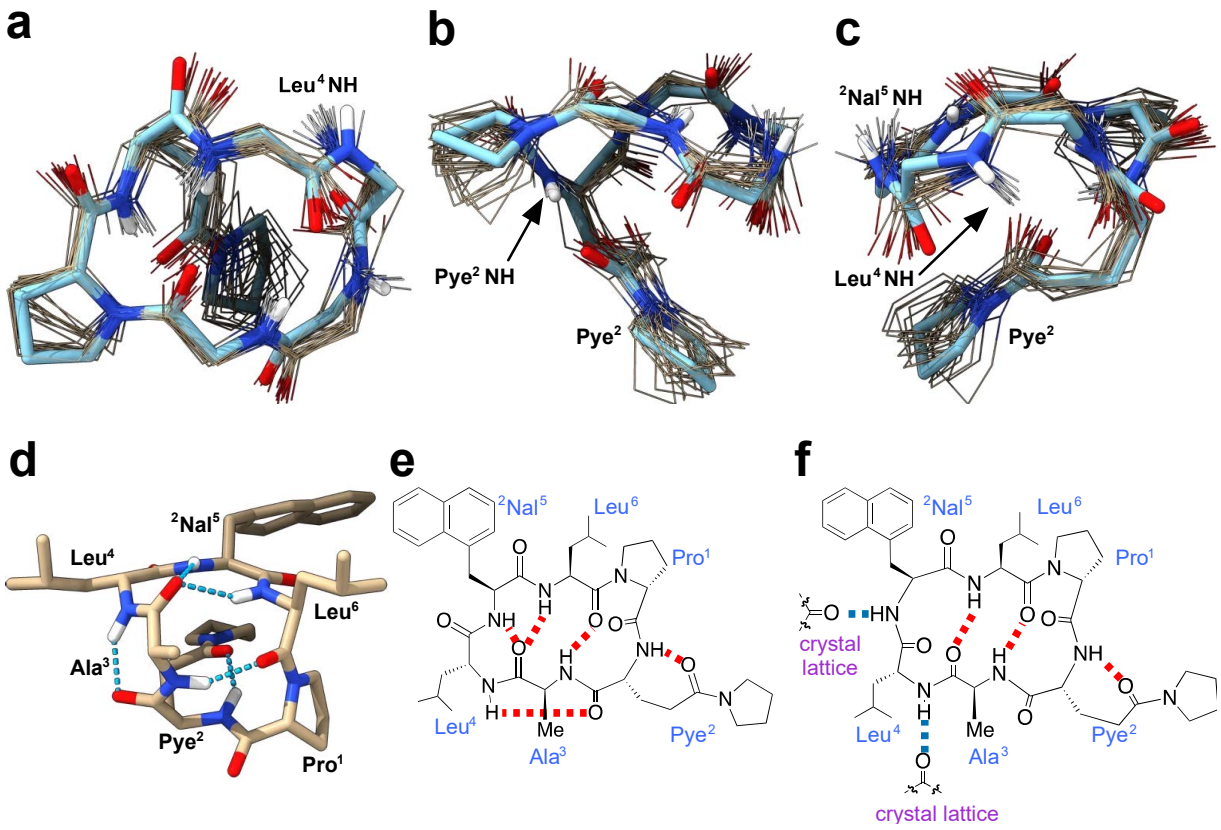
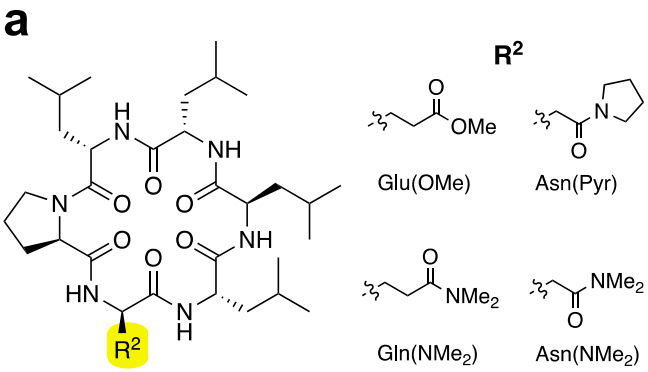


Figure 6. (**a-c**) Overlays of **3-Pye^2(Ala^{3}Nal^5)** crystal structure (blue) and 20 simulated McMD structures (gold) derived from NOESY information. Only the backbone and Pye side chain are shown for clarity. (**d**) Structure of simulated **3-Pye^2(Ala^{3}Nal^5)** having 5 IMHBs. (**e**) IMHB pattern of (**d**). (**f**) IMHB pattern of **3-Pye^2(Ala^{3}Nal^5)** crystallographic structure.

Effect of Different HBAs on Lipophilicity and Permeability

Data from various experiments above confirmed the lipophilicity and permeability improvements are due to the hydrogen bond from side chain HBA to the exposed HBD. To further expand the scope of HBAs that can be employed, we replaced Pye2 with four other HBAs, including *N,N*-dimethylamine substituent {**3-Gln(NMe$_2$)^2**}, methyl ester functional group {**3-Glu(OMe)^2**} and two short side chain amides {**3-Asn(Pyr)^2** and **3-Asn(NMe$_2$)^2**} (Figure 7a). Amide temperature coefficient experiments (Figure 7b) showed that changing substituent from pyrrolidine to *N,N*-dimethylamine {**3-Pye^2** to **3-Gln(NMe$_2$)^2**} did not significantly change the chemical shift or the temperature coefficient of its backbone NH. In contrast, replacing the amide side chain with an ester {**3-Pye^2** to **3-Glu(OMe)^2**} caused a significant upfield shift and much more negative $\Delta\delta_{NH}/\Delta T$ (-12.8 ppb/K), consistent with the decreased electron density at the ester carbonyl and the enthalpically weaker nature of the NH-to-ester hydrogen bond.

Shortening the Pye side chain by one methylene group {**3-Pye^2** to **3-Asn(Pyr)^2** or **3-Asn(NMe$_2$)^2**} shifted the backbone NH upfield and resulted in a much smaller (less negative) $\Delta\delta_{NH}/\Delta T$. These shifts relative to the longer-chain Gln derivatives is consistent with the pattern observed in the simple model compounds, with the decreased temperature shift resulting from the lower entropic penalty associated with the shorter Asn chain and the upfield shift resulting from the differences in the H-bond geometry between the two sets. However, McMD simulations of **3-Asn(Pyr)^2** in $CHCl_3$ offer another explanation: The shorter Asn side chain cannot reach its own backbone amide, although its close proximity to the NH groups of both residues 4 and 5 result in a bidentate H-bond with these two HBDs. Although this results in a completely new SC-BB H-bonding pattern, the overall backbone geometry is similar to that of **3-Pye^2** (Figure S12). The McMD ensemble of **3-Asn(Pyr)^2** shows no hydrogen bond to the amide NH of residue 2, consistent with its upfield shift and small temperature shift compared to the Gln derivatives (Figure S11). Testing these hypotheses will require further structural analysis.



ID	ALogP	log $D_{dec/w}$	LPE	P_{app}	solubility
3-Pye^2	2.46	1.1	3.9	1.5	730
3-Glu(OMe)^2	2.44	0.69	3.6	2.0	710
3-Gln(NMe$_2$)^2	2.00	0.43	3.8	0.70	1660
3-Asn(Pyr)^2	2.13	0.53	3.7	0.56	670
3-Asn(NMe$_2$)^2	1.67	0.24	3.9	0.28	1120

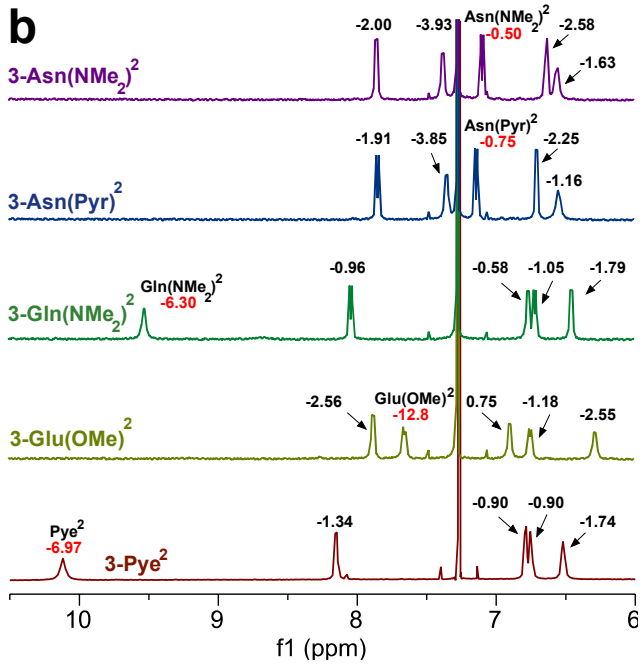


Figure 7. (**a**) Compound 3 derivatives bearing different HBAs on 2nd position and their experimental data. P_{app} x 10^{-6} cm/s, solubility in µM. (**b**) Amide-proton NMR spectra of compounds shown in (**a**). Amide NHs corresponding to the side-chain HBA residues are labeled, and the values indicate the $\Delta\delta_{NH}/\Delta T$ in ppb/K.

Lipophilicity and permeability of all four chemotypes were similar to the **3-Pye^2** scaffold (Figure 7a). LPEs of amide HBAs were within 0.2 log unit difference, while the LPE of the ester was slightly lower. Moreover, all compounds had PAMPA permeabilities within the same trend of **3-Pye^2** library (Figure S2). Notably, *N*,*N*-dimethylamine

substituent improved cyclic hexapeptide's solubility as expected from higher intrinsic hydrophilicity than pyrrolidine {**3-Pye**^{2} vs **3-Gln(NMe$_2$)**^{2} and **3-Asn(Pyr)**^{2} vs **3-Asn(NMe$_2$)**^{2}} More surprisingly, considering amides with same substituents {**3-Pye**^{2} vs **3-Asn(Pyr)**^{2} and **3-Gln(NMe$_2$)**^{2} vs **3-Asn(NMe$_2$)**^{2} }, cyclic hexapeptides bearing Gln side chain, despite having higher ALogP, were more soluble than Asn side chain. A short Asn side chain may form IMHB even in the polar solvent unlike a flexible Gln side chain that may have an equilibrium between IMHB and solvent-exposed formations. Alternatively, the alternate low-dielectric conformation calculated for **3-Asn(Pyr)**^{2} may be slightly less chameleonic, i.e., less able to access more polar conformations in aqueous solution. These experimental results broaden the opportunity to fine-tune lipophilicity of cyclic hexapeptide, for instance, by introducing a small amide substituent to enhance solubility of a highly hydrophobic cyclic hexapeptide, or a large amide substituent to slightly ameliorate solubility penalty while preserving overall permeability.

Conclusion

The inability of traditional small molecules to target large protein interfaces steers medicinal chemists to search for novel scaffolds in the bRo5 chemical space. Cyclic peptides have the potential to bind previously "undruggable" protein targets, but poor membrane permeability due to their bRo5 and peptidic nature has impeded the application toward intracellular targets. In this work, we introduce a potentially general approach to mask the amide HBD using SC-BB from simple *N*,*N*-dialkylated derivatives of Gln and Asn. Our studies in cyclic hexapeptide models showed that *N*,*N*-disubstituted Gln derivatives can serve as a side chain HBA to sequester the exposed backbone NH group in scaffold **1** and **3** when these units are substituted next to the proline on the C-terminus. Moreover, the SC-BB hydrogen bonding helps promote the IMHB network within the cyclic hexapeptides by restricting backbone conformation based on the McMD calculations. As a result, we observed an improvement in permeability in both artificial and cell-based assays. This method can be applied to any side chain combinations, and the HBA can be modified extensively by varying the functional group, amide substituents, or length of side chain. While polar and/or charged residues are often used to improve solubility in peptides, every polar proteinogenic residue also contains HBD, which are intrinsically deleterious to passive permeability. As a polar, non-proteinogenic residue that contains a strong HBA moiety but that lacks HBD, the Pye residue may allow medicinal chemists to modulate the chemical properties of cyclic peptides to balance their passive permeabilities, solubilities in the context of maintaining favorable target binding.

Masking exposed HBD by SC-BB hydrogen bonding could be used in conjunction with other methods to improve physico-chemical properties cyclic peptides. HBA-containing side chains can be introduced as pro-drugs, for

example, masking both the side chain functional group as well as the backbone amide in cases where SC-BB interactions are favorable.[65-67] However, since the Pye residue itself dramatically improves solubility while only slightly diminishing LPE, it may be generally useful for improving the properties of lipophilic cyclic peptide leads even in the absence of favorable SC-BB interactions. Therefore, this approach opens an opportunity to deliver cyclic peptides containing polar amino acids and backbone amide NH groups through the passive permeability pathway. Future studies will be directed toward evaluating the scope of Pye and similar residues in the context of improving the properties of bioactive compounds without abrogating their biochemical potency.

Experimental Section

All reactions were carried out under an inert atmosphere (nitrogen, or argon where stated) with dry solvents under anhydrous conditions. Glassware for anhydrous reactions was dried in an oven at 140 ˚C for minimum 6 h prior to use. Dry solvents were obtained by passing the previously degassed solvents through activated alumina columns. Reagents were purchased at a high commercial quality (typically 97 % or higher) and used without further purification, unless otherwise stated. Amino acids and amines were purchased from Combi- Blocks, Oakwood, Sigma-Aldrich, or Chem-Impex. COMU and HATU were purchased from Combi-Blocks or Chem-Impex. Piperidine was purchased from Spectrum Chemical. 1,9-Decadiene was purchased from TCI Chemicals. 2-Chlorotrityl chloride polystyrene resin was purchased from Rapp-Polymere. High field NMR spectra were recorded with Bruker Avance III HD at 500 MHz for ^{1}H and 126 MHz for ^{13}C or Bruker Ascend^{TM} with cryoprobe at 800 MHz for ^{1}H and 201 MHz for ^{13}C. NMR spectra were calibrated using residual non-deuterated solvent as an internal reference ($CDCl_3$: ^{1}H NMR = 7.27, ^{13}C NMR = 77.0). The following abbreviations were used to explain the multiplicities: s = singlet, d = doublet, t = triplet, q = quartet, quint = quintet, dd = double doublet, dt = double triplet, dq = double quartet, m = multiplet, br = broad. Purity of synthesized compounds were determined on an Advion AVANT HPLC-express**ion**® CMS system using a C18 Kinetex® colum (30 x 2.1 mm, 2.6 µm 100 Å) at a flow rate of 0.4 mL/min. The mobile phase was composed of 0.1% (v/v) formic acid in Milli-Q water (solvent A) and 0.1% formic acid in CH_3CN (solvent B). The gradient elution was ramped from 20 to 100% B over 6 minutes and held at 100% B for 1 minute. The detection was performed at 200 nm and 254 nm and the column temperature in an oven was 50 ºC. All individual synthesized compounds are ≥ 95% pure by HPLC or NMR analysis. **3-Pye^{2}** and **3-Pye^{3}** libraries were synthesized as purified mixtures and quantified via UPLC-MS and selected-mass monitoring.

Loading Pye on 2-Chlorotrityl Resin. 2-Chlorotrityl chloride resin was swelled in dried CH_2Cl_2 for 45 minutes. The solution of Pye and 2 eq of DIPEA in CH_2Cl_2 (0.16 M) was prepared 10 minutes prior adding to the resin. The resin was shaken in the solution for 3 h and drained. The mixture of 1:2:17 $^{i}Pr_2NEt$:MeOH:CH_2Cl_2 was added to cap any remaining active site. Resin was washed with CH_2Cl_2 (3x), DMF (3x) and CH_2Cl_2 (3x) and dried under vacuum overnight. The loading capacity was quantified by UV absorbance of the dibenzofulvene byproduct at 301 nm in ethanol by microcleavage of weighted resin.

Automatic Linear Peptide Synthesis. Linear peptides were synthesized using an automated peptide synthesizer (Prelude X, Protein Technologies). Commercial pre-loaded 2-chlorotrityl resin was used in 0.05 mmol scale. Fmoc deprotection was carried out with 20% piperidine in DMF for 3 minutes at 90 °C twice. Couplings were performed

using Fmoc-protected amino acids (4 eq), HATU (3.8 eq), and $^{i}Pr_2NEt$ (6 eq) in DMF (0.4 M with respect to amino acid) for 10 minutes at 90 °C. A capping step was performed after each amide coupling with a 1:1 mixture of acetic anhydride and $^{i}Pr_2NEt$ in DMF. Each coupling, deprotection, and capping step was followed by a wash with DMF (6x), CH_2Cl_2 (3x) and DMF (6x). Complete linear peptides were cleaved off resin with 30% HFIP in CH_2Cl_2 for 1 h three times with a CH_2Cl_2 wash equivalent to 5 resin volumes in between each step. Solvent was removed with Biotage V10 evaporator to obtain linear peptides as solid.

Peptide Cyclization. Crude linear peptides were dissolved in dry THF with 4 eq of $^{i}Pr_2NEt$ and added dropwise to a solution of 1:1 THF/MeCN containing 2 eq of COMU for a final concentration of 2 mM under argon atmosphere. Reactions were stirred for 12-24 h until complete cyclization was achieved as monitored by LC-MS. The reaction was reduced in vacuo and purified via reverse-phased column chromatography using a Biotage Isolera Prime.

Purification of Cyclic Hexapeptides. Purification of cyclic peptides were carried by Biotage Isolera Prime equipped with Biotage Sfär Bio C18 D 25 g column eluting with the 20-100% MeCN/H_2O gradient modified with 0.1% TFA at a flow rate of 40 mL/min.

Synthesis for 3-Pye^2 and 3-Pye^3 Libraries. Cyclic hexapeptide libraries were synthesize manually by split-pool approach (Scheme S3 and S4). Linear peptide sequences were synthesized from commercial pre-loaded 2-chlorotrityl-L-leucine resin starting with 200 mg for each R^1 amino acids (1 g total; 1.13 mmol). Fmoc deprotections were carried out with 20% piperidine in DMF for 20 minutes twice. Couplings were performed using Fmoc-protected amino acids (4 eq), HATU (4 eq), HOAt (4 eq) and $^{i}Pr_2NEt$ (8 eq) in DMF (0.3 M with respect to amino acid) for 1 h. After each coupling and deprotection step, the resin was washed with DMF (5x), CH_2Cl_2 (3x) and DMF (5x), and monitored the completeness of reaction by Kaiser test with small amount of resin. After coupling the R^2 amino acids, the resin was kept separately as the sub library. Linear peptides were cleaved in 30% HFIP / CH_2Cl_2 solution. Cyclization was performed same as mentioned above.

Synthesis of Pye peptoid monomer. 2-Chlorotrityl chloride resin was swelled in CH_2Cl_2 1 h before charging with 1M bromoacetic acid, 1M DIPEA in CH_2Cl_2. Resin was shaken for 1 h, then rinsed with DMF (3x) and CH_2Cl_2 (3x). 3-Amino-1-(pyrrolidin-1-yl)propan-1-one (4 eq) was first free-base by shaking in NMP in the presence of powdered KOH (12 eq). The KOH powder was separated by centrifuging at 16,000 x g for 5 minutes. Resin was shaken in this supernatant for 2 h and washed with DMF (3x) and CH_2Cl_2 (3x). Phenylacetic acid (3 eq), COMU (2.8 eq) and DIPEA (6 eq) were added to cap the resin. After successively wash with DMF (3x) and CH_2Cl_2 (3x), 3M HCl

generated by mixing acetyl chloride in anhydrous MeOH.[68] was added and the resin was shaken overnight to cleave the peptoid off the resin as well as undergoing methyl esterification simultaneously. The crude was dried under nitrogen stream and used for the experiment without further purification (scheme S5).

General Protocol of UPLC-MS Analysis. UPLC-MS analyses of cyclic peptides were performed via a Thermo Scientific Ultimate 3000 UPLC system, using a Thermo Hypersil GOLD C18 30 x 2.1 (1.9u) column (#25002-032130). Flow rate was set to 1 mL/min and a gradient method was as followed: 0.0-0.5 min, 5% MeCN; 0.5-0.75 min, ramp to 95% MeCN; 0.75-3.0 min, 95% MeCN; 3.0-3.5 min, 5% MeCN. Mass identification and quantification used an inline Thermo Scientific Orbitrap VelosPro (FTMS mode), tuned for maximum ionization of cyclosporin A, background ion locking on octyl phthalate, 200-2000 AMU mass windows, using +/- 0.02 AMU windows for integration.

Shake Flask Partition Experiment. Test compounds (400 μM, 8 μL) in DMSO were mixed in the solution of 800 μL 1,9-deacdiene and 800 μL PBS buffer in Eppendorf tubes and agitated by vortex (30 min) and sonication (30 min). Tubes were centrifuged for 10 min at 16,000 x g to separate two layers. 150 μL of each layer was transferred to the 96-well plate in quadruplicate and evaporated overnight in a Genevac centrifugal evaporator (60 °C). Test compounds were resuspended in 150 μL of 1:1 MeCN/H_2O and analyzed by UPLC-MS as described above.

LPE calculation: LPE was calculated from the equation published previously.[53]

$$\text{LPE} = \log D_{\text{dec/w}} - 1.06\text{ALogP} + 5.47$$

ALogP is a 2D molecular descriptor to represent the octanol/water partition coefficient determined from atoms in a molecule, which was calculated from Discovery Studio software.

PAMPA Assays: A 96-well donor plate with 0.45 μm hydrophobic Immobilon-P membrane supports (Millipore MAIPNTR10) was loaded with 5 μL of 1% lecithin in *n*-dodecane. Cyclic peptides in PBS solution containing 5% DMSO were loaded into donor plate (150 μL) and attached to the acceptor plate having 300 μL of 5% DMSO in PBS buffer. Each sample was run in quadruplicate for ~15 h at 20 ºC. The concentration of each compound in the donor and acceptor wells was quantified by UPLC-MS to calculate P_{app}.

Thermodynamic Solubility Assay: 20 μL of 10 mM stock solutions were dispensed into a 96-well conical plate and evaporated overnight in a Genevac centrifugal evaporator (60 °C). PBS (100 μL) was reintroduced to the plate to yield a maximum 2 mM concentration, then the plate was sealed and sonicated for 1 h. The plate was then gently

agitated at 37 °C for ~24 h. The mixtures were filtered through a 0.7 μm glass fiber filter plate (Agilent 200965-100) into a 96-well conical plate. The filtrate was further diluted up to 40 fold with MeCN in a new 96-well plate for quantification via UPLC-MS. Standard curves of each compound were acquired from serial dilution of stock solution with DMSO (50 μM to 0.1 μM) and used to calculate concentrations of analytes. All standards and analytes were performed in triplicate and averaged.

Amide Proton Temperature Coefficient Experiment: Approximately 3 mg of model compounds or cyclic peptides samples were dissolved in 550 μL $CDCl_3$ and ^1H NMR spectra were recorded at temperatures 300, 305, 310, 315, 320 and 323 K. Chloroform residue was used for calibration, and temperature coefficients of amide protons were calculated from best fit of experimental data points.

Caco-2 Cell Permeability: The Caco-2 assay was performed by Axcelead Drug Discovery Partners, Inc. (Kanagawa, Japan).

Conformational Search McMD Simulations: The procedure of the conformational search was same as previous reports.[52] Briefly, flat potential energy was obtained corresponding to the temperature between at T = 280 and 1505 K, and after production run, resampling method was used to obtain the ensemble at T = 300 K. From this ensemble, 5,000 conformers were used to further analysis.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Syntheses of model compounds **4** and amino acid building blocks; Details of solid-phase synthesis, *In silico* SC-BB hydrogen bonding, shake flask partition experiment, PAMPA assay, thermodynamic solubility assay, amide proton temperature coefficient experiment, NOE buildups and interproton distances, IMHB patterns, crystal structure of **3-Pye^2(Ala^{3}Nal^5)** and **3-Pye^2(HPhe^5)** (PDF)

The 20 structures of **3-Pye^2(Ala^{3}Nal^5)** from the McMD ensemble at T = 300K with the fewest violations from the NOESY-derived distances (PDB)

CCDC 2122986-2122987 contain the supplementary crystallographic data for this paper. These data can be obtained free of charge via www.ccdc.cam.ac.uk/data_request/cif, or by emailing data_request@ccdc.cam.ac.uk, or by contacting The Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK; fax: +44 1223 336033

AUTHOR INFORMATION

Corresponding Author

R. Scott Lokey – Department of Chemistry and Biochemistry, University of California, Santa Cruz, USA; Email: slokey@ucsc.edu

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

This research was supported by funding from the National Institutes of Health (GM131135). X-ray diffraction studies were performed on an instrument purchased with NSF MRI grant #2018501. We also thank Prof. Sam Gellman for helpful suggestions during the preparation of this manuscript.

ABBREVIATIONS USED

AA, arbitrary amino acid; ALogP, atomistic calculated octanol/water partition coefficient; Boc_2O, di-tert-butyl decarbonate; bRo5, beyond-Rule of 5; Caco-2, human colorectal adenocarcinoma cells; COMU, (1-cyano-2-ethoxy-2-oxoethylidenaminooxy)dimethylaminomorpholino-carbenium hexafluorophosphate; DBU, 1,8-diazabicyclo[5.4.0]undec-7-ene; DCM, dichloromethane; DMAP, 4-dimethylaminopyridine; DMF, *N*,*N*-dimethylformamide; DMSO, dimethylsulfoxide; EDC, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide; Fmoc, 9-fluorenylmethoxycarbonyl; HATU, 1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate; HB, hydrogen bond; HBA, hydrogen bond acceptor; HBD, hydrogen bond donor; HFIP, 1,1,1,3,3,3-hexafluoro-2-propanol; HOBt, hydroxybenzotriazole; IMHB, intramolecular hydrogen bond; log *D*, shake flask distribution coefficient at pH 7.4; LPE, Lipophilic Permeability Efficiency; MeCN, acetonitrile; MeOH, methanol; MW, molecular weight; NMM, *N*-methylmorpholine; PAMPA; cell-free parallel artificial membrane permeability assay; P_{app}, apparent permeation rate (A-to-B, $x10^{-6}$ cm/s); PBS, phosphate-buffered saline; Ro5, Rule of 5; SPPS, solid phase peptide synthesis; TFA, trifluoroacetic acid; THF, tetrahydrofuran; UPLC-MS, ultra-high performance liquid chromatography mass spectrometry;

References

(1) Hewitt, W. M.; Leung, S. S.; Pye, C. R.; Ponkey, A. R.; Bednarek, M.; Jacobson, M. P.; Lokey, R. S. Cell-Permeable Cyclic Peptides from Synthetic Libraries Inspired by Natural Products. *J. Am. Chem. Soc.* **2015,** *137*, 715-721.

(2) Shiroma, Y.; Takahashi, R.-u.; Yamamoto, Y.; Tahara, H. Targeting DNA Binding Proteins for Cancer Therapy. *Cancer Sci.* **2020,** *111*, 1058-1064.

(3) Wu, P. Inhibition of RNA-Binding Proteins with Small Molecules. *Nat. Rev. Chem.* **2020,** *4*, 441-458.

(4) Hong, S. RNA Binding Protein as an Emerging Therapeutic Target for Cancer Prevention and Treatment. *J. Cancer Prev.* **2017,** *22*, 203-210.

(5) Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L. How Many Drug Targets Are There? *Nat. Rev. Drug Discov.* **2006,** *5*, 993-996.

(6) Lu, H.; Zhou, Q.; He, J.; Jiang, Z.; Peng, C.; Tong, R.; Shi, J. Recent Advances in the Development of Protein–Protein Interactions Modulators: Mechanisms and Clinical Trials. *Signal Transduct. Target. Ther.* **2020,** *5*, 213.

(7) Naylor, M. R.; Bockus, A. T.; Blanco, M.-J.; Lokey, R. S. Cyclic Peptide Natural Products Chart the Frontier of Oral Bioavailability in the Pursuit of Undruggable Targets. *Curr. Opin. Chem. Biol.* **2017,** *38*, 141-147.

(8) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Deliv. Rev.* **1997,** *23*, 3-25.

(9) Li, Y.; Li, W.; Xu, Z. Improvement on Permeability of Cyclic Peptide/Peptidomimetic: Backbone N-Methylation as a Useful Tool. *Mar. Drugs* **2021,** *19*, 311.

(10) Villar, E. A.; Beglov, D.; Chennamadhavuni, S.; Porco, J. A.; Kozakov, D.; Vajda, S.; Whitty, A. How Proteins Bind Macrocycles. *Nat. Chem. Biol.* **2014,** *10*, 723-731.

(11) Doak, B. C.; Zheng, J.; Dobritzsch, D.; Kihlberg, J. How Beyond Rule of 5 Drugs and Clinical Candidates Bind to Their Targets. *J. Med. Chem.* **2016,** *59*, 2312-2327.

(12) Doak, Bradley C.; Over, B.; Giordanetto, F.; Kihlberg, J. Oral Druggable Space Beyond the Rule of 5: Insights from Drugs and Clinical Candidates. *Chem. Biol.* **2014,** *21*, 1115-1142.

(13) Caron, G.; Kihlberg, J.; Goetz, G.; Ratkova, E.; Poongavanam, V.; Ermondi, G. Steering New Drug Discovery Campaigns: Permeability, Solubility, and Physicochemical Properties in the BRo5 Chemical Space. *ACS Med. Chem. Lett.* **2021,** *12*, 13-23.

(14) Shultz, M. D. Two Decades under the Influence of the Rule of Five and the Changing Properties of Approved Oral Drugs. *J. Med. Chem.* **2019,** *62*, 1701-1714.

(15) Brown, D. G.; Wobst, H. J. A Decade of FDA-Approved Drugs (2010–2019): Trends and Future Directions. *J. Med. Chem.* **2021,** *64*, 2312-2338.

(16) Jing, X.; Jin, K. A Gold Mine for Drug Discovery: Strategies to Develop Cyclic Peptides into Therapies. *Med. Res. Rev.* **2020,** *40*, 753-810.

(17) Dougherty, P. G.; Sahni, A.; Pei, D. Understanding Cell Penetration of Cyclic Peptides. *Chem. Rev.* **2019,** *119*, 10241-10287.

(18) Buckton, L. K.; Rahimi, M. N.; McAlpine, S. R. Cyclic Peptides as Drugs for Intracellular Targets: The Next Frontier in Peptide Therapeutic Development. *Chem. Eur. J.* **2021,** *27*, 1487-1513.

(19) Nielsen, D. S.; Shepherd, N. E.; Xu, W.; Lucke, A. J.; Stoermer, M. J.; Fairlie, D. P. Orally Absorbed Cyclic Peptides. *Chem. Rev.* **2017,** *117*, 8094-8128.

(20) Guillen Schlippe, Y. V.; Hartman, M. C. T.; Josephson, K.; Szostak, J. W. In Vitro Selection of Highly Modified Cyclic Peptides That Act as Tight Binding Inhibitors. *J. Am. Chem. Soc.* **2012,** *134*, 10469-10477.

(21) Witek, J.; Keller, B. G.; Blatter, M.; Meissner, A.; Wagner, T.; Riniker, S. Kinetic Models of Cyclosporin A in Polar and Apolar Environments Reveal Multiple Congruent Conformational States. *J. Chem. Inf. Model.* **2016,** *56*, 1547-1562.

(22) Wang, C. K.; Swedberg, J. E.; Harvey, P. J.; Kaas, Q.; Craik, D. J. Conformational Flexibility Is a Determinant of Permeability for Cyclosporin. *J. Phys. Chem. B* **2018,** *122*, 2261-2276.

(23) Loosli, H.-R.; Kessler, H.; Oschkinat, H.; Weber, H.-P.; Petcher, T. J.; Widmer, A. Peptide Conformations. Part 31. The Conformation of Cyclosporin A in the Crystal and in Solution. *Helv. Chim. Acta* **1985,** *68*, 682-704.

(24) Caron, G.; Kihlberg, J.; Ermondi, G. Intramolecular Hydrogen Bonding: An Opportunity for Improved Design in Medicinal Chemistry. *Med. Res. Rev.* **2019,** *39*, 1707-1729.

(25) Furukawa, A.; Townsend, C. E.; Schwochert, J.; Pye, C. R.; Bednarek, M. A.; Lokey, R. S. Passive Membrane Permeability in Cyclic Peptomer Scaffolds Is Robust to Extensive Variation in Side Chain Functionality and Backbone Geometry. *J. Med. Chem.* **2016,** *59*, 9503-9512.

(26) Wang, C. K.; Northfield, S. E.; Colless, B.; Chaousis, S.; Hamernig, I.; Lohman, R.-J.; Nielsen, D. S.; Schroeder, C. I.; Liras, S.; Price, D. A.; Fairlie, D. P.; Craik, D. J. Rational Design and Synthesis of an Orally Bioavailable Peptide Guided by NMR Amide Temperature Coefficients. *Proc. Natl. Acad. Sci. U. S. A.* **2014,** *111*, 17504-17509.

(27) Wang, C. K.; Northfield, S. E.; Swedberg, J. E.; Colless, B.; Chaousis, S.; Price, D. A.; Liras, S.; Craik, D. J. Exploring Experimental and Computational Markers of Cyclic Peptides: Charting Islands of Permeability. *Eur. J. Med. Chem.* **2015,** *97*, 202-213.

(28) Schwochert, J.; Lao, Y.; Pye, C. R.; Naylor, M. R.; Desai, P. V.; Gonzalez Valcarcel, I. C.; Barrett, J. A.; Sawada, G.; Blanco, M. J.; Lokey, R. S. Stereochemistry Balances Cell Permeability and Solubility in the Naturally Derived Phepropeptin Cyclic Peptides. *ACS Med. Chem. Lett.* **2016,** *7*, 757-761.

(29) Biron, E.; Chatterjee, J.; Ovadia, O.; Langenegger, D.; Brueggen, J.; Hoyer, D.; Schmid, H. A.; Jelinek, R.; Gilon, C.; Hoffman, A.; Kessler, H. Improving Oral Bioavailability of Peptides by Multiple N-Methylation: Somatostatin Analogues. *Angew. Chem. Int. Ed.* **2008,** *47*, 2595-2599.

(30) Bockus, A. T.; Schwochert, J. A.; Pye, C. R.; Townsend, C. E.; Sok, V.; Bednarek, M. A.; Lokey, R. S. Going out on a Limb: Delineating the Effects of β-Branching, N-Methylation, and Side Chain Size on the Passive Permeability, Solubility, and Flexibility of Sanguinamide a Analogues. *J. Med. Chem.* **2015,** *58*, 7409-7418.

(31) Kwon, Y.-U.; Kodadek, T. Quantitative Comparison of the Relative Cell Permeability of Cyclic and Linear Peptides. *Chem. Biol.* **2007,** *14*, 671-677.

(32) Tan, N. C.; Yu, P.; Kwon, Y.-U.; Kodadek, T. High-Throughput Evaluation of Relative Cell Permeability between Peptoids and Peptides. *Bioorg. Med. Chem.* **2008,** *16*, 5853-5861.

(33) Räder, A. F. B.; Reichart, F.; Weinmüller, M.; Kessler, H. Improving Oral Bioavailability of Cyclic Peptides by N-Methylation. *Bioorg. Med. Chem.* **2018,** *26*, 2766-2773.

(34) Thansandote, P.; Harris, R. M.; Dexter, H. L.; Simpson, G. L.; Pal, S.; Upton, R. J.; Valko, K. Improving the Passive Permeability of Macrocyclic Peptides: Balancing Permeability with Other Physicochemical Properties. *Bioorg. Med. Chem.* **2015,** *23*, 322-327.

(35) Hill, T. A.; Lohman, R.-J.; Hoang, H. N.; Nielsen, D. S.; Scully, C. C. G.; Kok, W. M.; Liu, L.; Lucke, A. J.; Stoermer, M. J.; Schroeder, C. I.; Chaousis, S.; Colless, B.; Bernhardt, P. V.; Edmonds, D. J.; Griffith, D. A.; Rotter, C. J.; Ruggeri, R. B.; Price, D. A.; Liras, S.; Craik, D. J.; Fairlie, D. P. Cyclic Penta- and Hexaleucine Peptides without N-Methylation Are Orally Absorbed. *ACS Med. Chem. Lett.* **2014,** *5*, 1148-1151.

(36) Nielsen, D. S.; Hoang, H. N.; Lohman, R.-J.; Hill, T. A.; Lucke, A. J.; Craik, D. J.; Edmonds, D. J.; Griffith, D. A.; Rotter, C. J.; Ruggeri, R. B.; Price, D. A.; Liras, S.; Fairlie, D. P. Improving on Nature: Making a Cyclic Heptapeptide Orally Bioavailable. *Angew. Chem. Int. Ed.* **2014,** *53*, 12059-12063.

(37) Vollbrecht, L.; Steinmetz, H.; Hofle, G.; Oberer, L.; Rihs, G.; Bovermann, G.; von Matt, P. Argyrins, Immunosuppressive Cyclic Peptides from Myxobacteria. II. Structure Elucidation and Stereochemistry. *J Antibiot (Tokyo)* **2002,** *55*, 715-721.

(38) Nyfeler, B.; Hoepfner, D.; Palestrant, D.; Kirby, C. A.; Whitehead, L.; Yu, R.; Deng, G.; Caughlan, R. E.; Woods, A. L.; Jones, A. K.; Barnes, S. W.; Walker, J. R.; Gaulis, S.; Hauy, E.; Brachmann, S. M.; Krastel, P.; Studer, C.; Riedl, R.; Estoppey, D.; Aust, T.; Movva, N. R.; Wang, Z.; Salcius, M.; Michaud, G. A.; McAllister, G.; Murphy, L. O.; Tallarico, J. A.; Wilson, C. J.; Dean, C. R. Identification of Elongation Factor G as the Conserved Cellular Target of Argyrin B. *PLOS ONE* **2012,** *7*, e42657.

(39) Zaretsky, S.; Hickey, J. L.; Tan, J.; Pichugin, D.; St. Denis, M. A.; Ler, S.; Chung, B. K. W.; Scully, C. C. G.; Yudin, A. K. Mechanistic Investigation of Aziridine Aldehyde-Driven Peptide Macrocyclization: The Imidoanhydride Pathway. *Chem. Sci.* **2015,** *6*, 5446-5455.

(40) Hickey, J. L.; Zaretsky, S.; St. Denis, M. A.; Kumar Chakka, S.; Morshed, M. M.; Scully, C. C. G.; Roughton, A. L.; Yudin, A. K. Passive Membrane Permeability of Macrocycles Can Be Controlled by Exocyclic Amide Bonds. *J. Med. Chem.* **2016,** *59*, 5368-5376.

(41) Zaretsky, S.; Scully, C. C. G.; Lough, A. J.; Yudin, A. K. Exocyclic Control of Turn Induction in Macrocyclic Peptide Scaffolds. *Chem. Eur. J.* **2013,** *19*, 17668-17672.

(42) Appavoo, S. D.; Huh, S.; Diaz, D. B.; Yudin, A. K. Conformational Control of Macrocycles by Remote Structural Modification. *Chem. Rev.* **2019,** *119*, 9724-9752.

(43) Driver, R. W.; Hoang, H. N.; Abbenante, G.; Fairlie, D. P. A Cyclic β-Strand Tripeptide with an α-Helix Like CD Spectrum. *Org. Lett.* **2009,** *11*, 3092-3095.

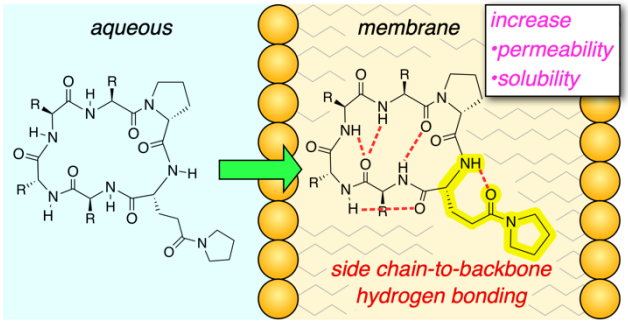
(44) Grotenbreg, G. M.; Timmer, M. S. M.; Llamas-Saiz, A. L.; Verdoes, M.; van der Marel, G. A.; van Raaij, M. J.; Overkleeft, H. S.; Overhand, M. An Unusual Reverse Turn Structure Adopted by a Furanoid Sugar Amino Acid Incorporated in Gramicidin S. *J. Am. Chem. Soc.* **2004,** *126*, 3444-3446.

(45) Wang, S.; König, G.; Roth, H.-J.; Fouché, M.; Rodde, S.; Riniker, S. Effect of Flexibility, Lipophilicity, and the Location of Polar Residues on the Passive Membrane Permeability of a Series of Cyclic Decapeptides. *J. Med. Chem.* **2021,** *64*, 12761-12773.

(46) Cierpicki, T.; Otlewski, J. Amide Proton Temperature Coefficients as Hydrogen Bond Indicators in Proteins. *J. Biomol. NMR* **2001,** *21*, 249-261.

(47) Baxter, N. J.; Williamson, M. P. Temperature Dependence of 1H Chemical Shifts in Proteins. *J. Biomol. NMR* **1997,** *9*, 359-369.
(48) Wagner, G.; Pardi, A.; Wuethrich, K. Hydrogen Bond Length and Proton NMR Chemical Shifts in Proteins. *J. Am. Chem. Soc.* **1983,** *105*, 5948-5949.
(49) Gellman, S. H.; Adams, B. R. Intramolecular Hydrogen Bonding in Simple Diamides. *Tetrahedron Lett.* **1989,** *30*, 3381-3384.
(50) Gellman, S. H.; Dado, G. P.; Liang, G. B.; Adams, B. R. Conformation-Directing Effects of a Single Intramolecular Amide-Amide Hydrogen Bond: Variable-Temperature NMR and IR Studies on a Homologous Diamide Series. *J. Am. Chem. Soc.* **1991,** *113*, 1164-1173.
(51) Andersen, N. H.; Neidigh, J. W.; Harris, S. M.; Lee, G. M.; Liu, Z.; Tong, H. Extracting Information from the Temperature Gradients of Polypeptide NH Chemical Shifts. 1. The Importance of Conformational Averaging. *J. Am. Chem. Soc.* **1997,** *119*, 8547-8561.
(52) Ono, S.; Naylor, M. R.; Townsend, C. E.; Okumura, C.; Okada, O.; Lokey, R. S. Conformation and Permeability: Cyclic Hexapeptide Diastereomers. *J. Chem. Inf. Model.* **2019,** *59*, 2952-2963.
(53) Naylor, M. R.; Ly, A. M.; Handford, M. J.; Ramos, D. P.; Pye, C. R.; Furukawa, A.; Klein, V. G.; Noland, R. P.; Edmondson, Q.; Turmon, A. C.; Hewitt, W. M.; Schwochert, J.; Townsend, C. E.; Kelly, C. N.; Blanco, M. J.; Lokey, R. S. Lipophilic Permeability Efficiency Reconciles the Opposing Roles of Lipophilicity in Membrane Permeability and Aqueous Solubility. *J. Med. Chem.* **2018,** *61*, 11169-11182.
(54) Ghose, A. K.; Crippen, G. M. Atomic Physicochemical Parameters for Three-Dimensional-Structure-Directed Quantitative Structure-Activity Relationships. 2. Modeling Dispersive and Hydrophobic Interactions. *J. Chem. Inf. Comput. Sci.* **1987,** *27*, 21-35.
(55) Thompson, S. J.; Hattotuwagama, C. K.; Holliday, J. D.; Flower, D. R. On the Hydrophobicity of Peptides: Comparing Empirical Predictions of Peptide Log P Values. *Bioinformation* **2006,** *1*, 237-241.
(56) Duban, M. E.; Bures, M. G.; DeLazzer, J.; Martin, Y. C. Virtual Screening of Molecular Properties: A Comparison of Log P Calculators. In *Pharmacokinetic Optimization in Drug Research*, 2001; pp 483-497.
(57) Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J. Prediction of Hydrophobic (Lipophilic) Properties of Small Organic Molecules Using Fragmental Methods: An Analysis of AlogP and ClogP Methods. *J. Phys. Chem. A* **1998,** *102*, 3762-3772.
(58) Kansy, M.; Senner, F.; Gubernator, K. Physicochemical High Throughput Screening: Parallel Artificial Membrane Permeation Assay in the Description of Passive Absorption Processes. *J. Med. Chem.* **1998,** *41*, 1007-1010.
(59) Fujikawa, M.; Nakao, K.; Shimizu, R.; Akamatsu, M. QSAR Study on Permeability of Hydrophobic Compounds with Artificial Membranes. *Bioorg. Med. Chem.* **2007,** *15*, 3756-3767.
(60) Sawada, G. A.; Barsuhn, C. L.; Lutzke, B. S.; Houghton, M. E.; Padbury, G. E.; Ho, N. F. H.; Raub, T. J. Increased Lipophilicity and Subsequent Cell Partitioning Decrease Passive Transcellular Diffusion of Novel, Highly Lipophilic Antioxidants. *J. Pharmacol. Exp. Ther.* **1999,** *288*, 1317-1326.
(61) Hu, H.; Krishnamurthy, K. Revisiting the Initial Rate Approximation in Kinetic Noe Measurements. *J. Magn. Reson.* **2006,** *182*, 173-177.
(62) Butts, C. P.; Jones, C. R.; Towers, E. C.; Flynn, J. L.; Appleby, L.; Barron, N. J. Interproton Distance Determinations by NOE – Surprising Accuracy and Precision in a Rigid Organic Molecule. *Org. Biomol. Chem.* **2011,** *9*, 177-184.
(63) Begnini, F.; Poongavanam, V.; Atilaw, Y.; Erdelyi, M.; Schiesser, S.; Kihlberg, J. Cell Permeability of Isomeric Macrocycles: Predictions and NMR Studies. *ACS Med. Chem. Lett.* **2021,** *12*, 983-990.
(64) Nakajima, N.; Nakamura, H.; Kidera, A. Multicanonical Ensemble Generated by Molecular Dynamics Simulation for Enhanced Conformational Sampling of Peptides. *J. Phys. Chem. B* **1997,** *101*, 817-824.
(65) Barlow, N.; Chalmers, D. K.; Williams-Noonan, B. J.; Thompson, P. E.; Norton, R. S. Improving Membrane Permeation in the Beyond Rule-of-Five Space by Using Prodrugs to Mask Hydrogen Bond Donors. *ACS Chem. Biol.* **2020,** *15*, 2070-2078.
(66) Simplício, A. L.; Clancy, J. M.; Gilmer, J. F. Prodrugs for Amines. *Molecules* **2008,** *13*, 519-547.
(67) Wang, D.; Zou, L.; Jin, Q.; Hou, J.; Ge, G.; Yang, L. Human Carboxylesterases: A Comprehensive Review. *Acta Pharm. Sin. B* **2018,** *8*, 699-712.
(68) Turner, R. A.; Weber, R. J.; Lokey, R. S. Direct Conversion of Resin-Bound Peptides to C-Terminal Esters. *Org. Lett.* **2010,** *12*, 1852-1855.

Table of Contents graphic



Supporting Materials

A New Amino Acid for Improving Permeability and Solubility in Macrocyclic Peptides through Side Chain-to-Backbone Hydrogen Bonding

Jaru Taechalertpaisarn[†], Satoshi Ono[‡], Okimasa Okada[‡], Timothy C. Johnstone[†], R. Scott Lokey*[,†]

[†]Department of Chemistry and Biochemistry, University of California, Santa Cruz, 1156 High Street, Santa Cruz, California 95064, United States

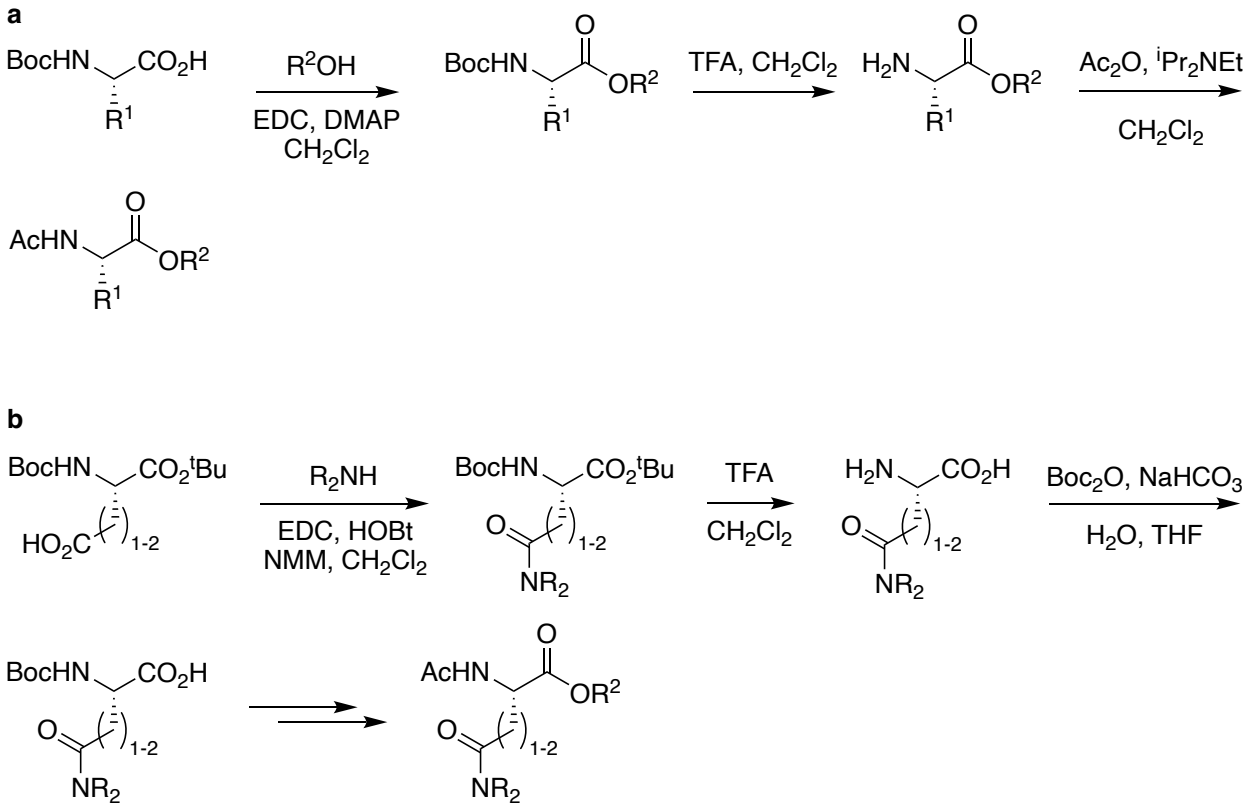
[‡]Modality Laboratories, Innovative Research Division, Mitsubishi Tanabe Pharma Corporation, 1000 Kamoshida-cho, Aoba-ku, Yokohama, Kanagawa 227-0033, Japan

*Correspondence to slokey@ucsc.edu

Table of Contents

Syntheses of Model Structures **4** S3
General Synthesis of Model Compounds (Scheme S1a) S3
General Synthesis of Substituted-amide Model Compounds (Scheme S1b) S4
General Synthesis of *N*-Fmoc-Amino Acids with Side-Chain Carboxamides. S4
Synthesis of 3-Amino-1-(pyrrolidin-1-yl)propan-1-one S11
Synthesis Scheme of Cyclic Peptides and Peptoid Monomer S13
In Silico Investigation of Side Chain-to-Backbone Hydrogen Bonding of Model Structures S15
Analytical Procedures S16
General Protocol of UPLC-MS Analysis S16
Shake Flask Partition Experiment S16
Parallel Artificial Membrane Permeability Assay (PAMPA) S16
Thermodynamic Solubility Assay S17
NMR Spectra of Amide Proton Temperature Coefficient Experiment S20
NOE Buildups and Interproton Distances S35
Intramolecular Hydrogen Bonding (IMHB) Patterns S41
Crystal Data and Structure Refinement for **3-Pye^{2}(HPhe^{5})** S44
Crystal Data and Structure Refinement for **3-Pye^{2}(Ala^{3}Nal^{5})** S46
LC-MS Spectra S49
References S62

Syntheses of Model Structures 4



Scheme S1. Syntheses of model structures **4** with (**a**) accessible amino acids (R^1 = Me, -$CH(CH_3)_2$, -$CH_2CH(CH_3)_2$, -$CH(CH_3)CH_2CH_3$, -$CH_2CH_2CONH_2$, -$CH_2CH_2CO_2Me$) and (**b**) substituted amides. R^2OH = 4-phenyl-1-butanol; R_2NH = pyrrolidine or dimethylamine

General Synthesis of Model Compounds (Scheme S1a)

Boc-protected amino acids (1.5 eq), DMAP (1.5 eq) and 4-phenyl-1-butanol (1 eq) were dissolved in CH_2Cl_2 (0.1 M) at 0 ºC. EDC (1.5 eq) was added subsequently, and the reaction was stirred under Ar atmosphere at ambient temperature. Solution was diluted in CH_2Cl_2 and extracted with 10% citric acid (3x), saturated $NaHCO_3$ (3x) and brine. Organic layer was dried over $MgSO_4$ and reduced to obtain the product as an oil, which was used in the next step without further purification.

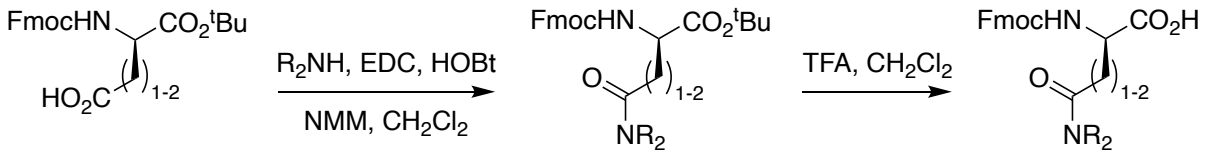
Boc-AA-OC_4H_8Ph (1 mmol) was dissolved in 3 mL CH_2Cl_2, then 3 mL of TFA was added and the reaction was stirred at room temperature for 3 h. Solution was completely dried, and the crude oil was used in the next step without further purification.

Crude material from previous step (0.5 mmol) was charged with 3 mL dry CH_2Cl_2 under Ar at 0 ºC. $^iPr_2NEt$ (2 eq) was added followed by Ac_2O (2 eq). The reaction was stirred overnight at ambient temperature. Organic solution was extracted with saturated NH_4Cl, $NaHCO_3$ and brine. After reducing the solution, product was purified by short column chromatography (50% EtOAc/hexanes then 5% MeOH/CH_2Cl_2) to obtain final product as oil.

General Synthesis of Substituted-amide Model Compounds (Scheme S1b)

Boc-Asp-O^{t}Bu or Boc-Glu-O^{t}Bu (1 mmol) and HOBt (1.5 eq) were dissolved in 10 mL CH_2Cl_2 at 0 ºC. EDC (1.5 eq) was added in one portion and the reaction was stirred at 0 ºC for 10 minutes followed by *N*-methylmorpholine (NMM, 3 eq) and secondary amine (pyrrolidine or dimethylamine, 1.5 eq). Reaction was allowed to stir at ambient temperature overnight under Ar. Solution was diluted in CH_2Cl_2 and extracted with 10% citric acid (3x), saturated $NaHCO_3$ (3x) and brine. Organic layer was dried over $MgSO_4$ and reduced to obtain product as solid. This solid compound was dissolved in 1:1 CH_2Cl_2/TFA solution to deprotect all acid-labile protecting groups. Once reaction was completed, solution was dried in vacuo, and the crude oil was azeotroped with diethyl ether and dried under vacuum overnight. This unprotected compound was then dissolved in 1:1 H_2O/THF and $NaHCO_3$ (2 eq) was added to the solution. Boc_2O (2 eq) was added, and the reaction was stirred overnight while pH was kept around 8. After reaction was completed, organic solvent was removed in vacuo, and the aqueous solution was extracted with diethyl ether (3x). The aqueous layer was acidified with citric acid until pH ~4 and extracted with EtOAc (5x). Organic layer was combined and reduced to obtain the Boc-protected amide-substituted amino acids. The remaining synthesis was done similar to the general synthesis described for Scheme S1a

General Synthesis of *N*-Fmoc-Amino Acids with Side-Chain Carboxamides.



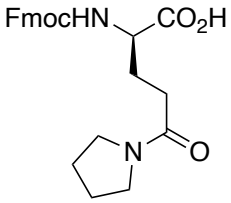
R_2NH = pyrrolidine, dimethylamine

Scheme S2. Synthesis scheme of Fmoc-amino acid building blocks.

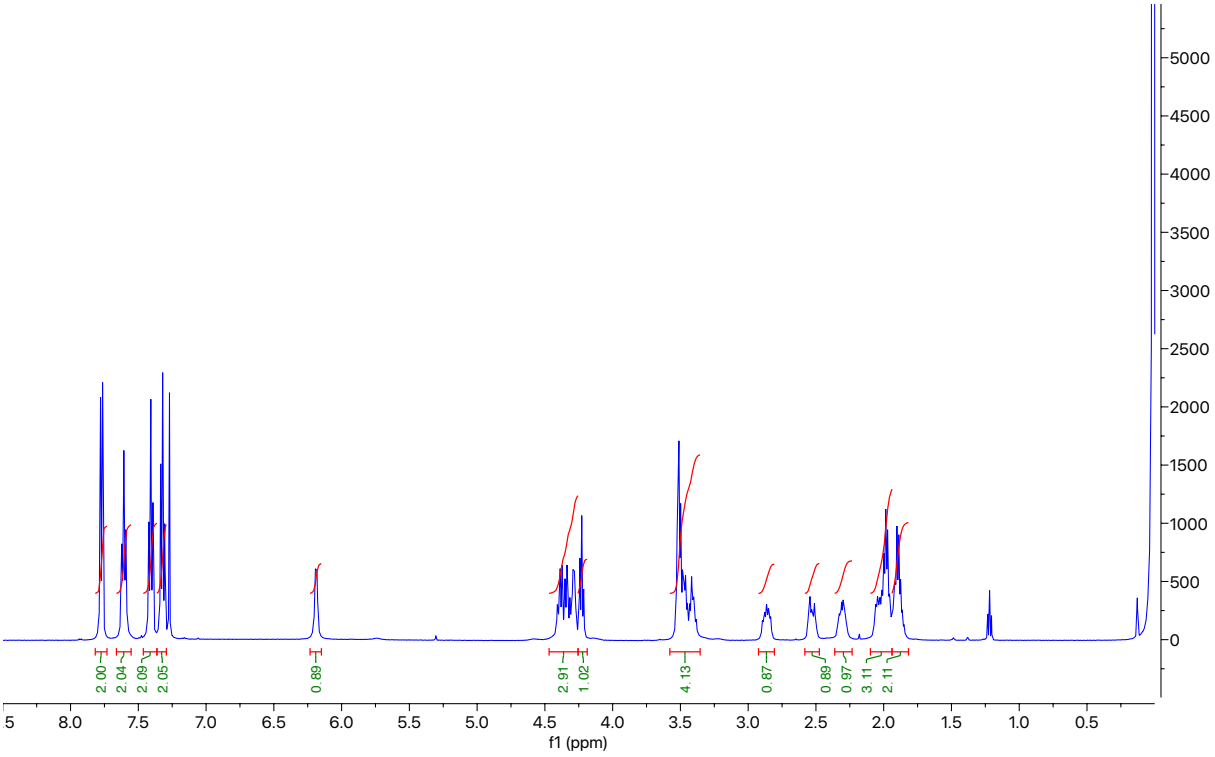
A detail synthesis of D-Fmoc-Pye-OH was shown as an example. D-Fmoc-Glu-OtBu (10 g, 23.5 mmol) and hydroxybenzotriazole (HOBt; 4.76 g, 35.3 mmol) were dissolved in 200 mL CH_2Cl_2 at 0 ºC. 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC; 6.76 g, 35.3 mmol) was added in one portion, and the reaction was stirred for 10 minutes. *N*-Methylmorpholine (NMM; 7.8 mL, 70.5 mmol) was added followed by pyrrolidine (2.3 mL, 28.2 mmol). The reaction was stirred under argon overnight at the ambient temperature. The solution was extracted 3 times with 10% citric acid, 3 times with saturated sodium bicarbonate solution and brine. The organic layer was dried and reduced. The resulting colorless oil was treated with a solution of 1:1 TFA/CH_2Cl_2 (50 mL) and stirred until the reaction was completed monitored by TLC. The solution was removed under reduced pressure and successively evaporated remaining TFA residue with CH_2Cl_2. Approximately 150 mL of diethyl ether was added to the crude and the solution was stirred vigorously overnight until the white precipitate was formed. The white solid was filtered and rinsed with diethyl ether to obtain the final product, D-Pye, as white solid (8.92 g, 90%)

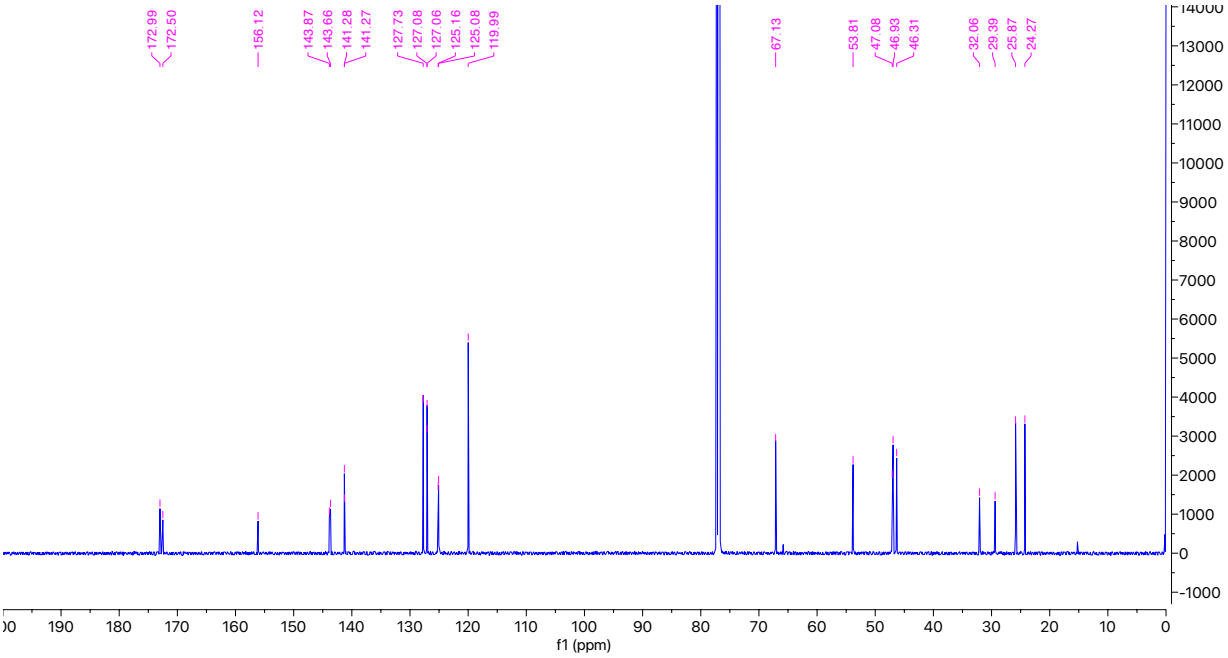
L-Fmoc-Pye-OH, D-Fmoc-Gln(NMe_2)-OH, D-Fmoc-Asn(Pyr)-OH and D-Fmoc-Asn(NMe_2)-OH were synthesized as described above.

(*R*)-2-((((9*H*-Fluoren-9-yl)methoxy)carbonyl)amino)-5-oxo-5-(pyrrolidin-1-yl)pentanoic acid (D-Fmoc-Pye-OH)

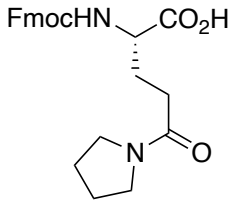


White solid, 90 % yield. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.77 (d, *J* = 7.5 Hz, 2H), 7.61 (t, *J* = 7.6 Hz, 2H), 7.41 (t, *J* = 7.5 Hz, 2H), 7.32 (t, *J* = 7.5 Hz, 2H), 6.19 (d, *J* = 5.9 Hz, 1H), 4.47 – 4.26 (m, 3H), 4.23 (t, *J* = 7.2 Hz, 1H), 3.58 – 3.35 (m, 4H), 2.86 (m, 1H), 2.53 (m, 1H), 2.36 – 2.23 (m, 1H), 2.10-1.94 (m, 3H), 1.94-1.82 (m, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 172.99, 172.50, 156.12, 143.87, 143.66, 141.28, 141.27, 127.73, 127.08, 127.06, 125.16, 125.08, 119.99, 67.13, 53.81, 47.08, 46.93, 46.31, 32.06, 29.39, 25.87, 24.27. HRMS (ESI) m/z calcd for $C_{24}H_{27}N_2O_5^+$ (M+H$^+$) 423.1914; found 423.1909

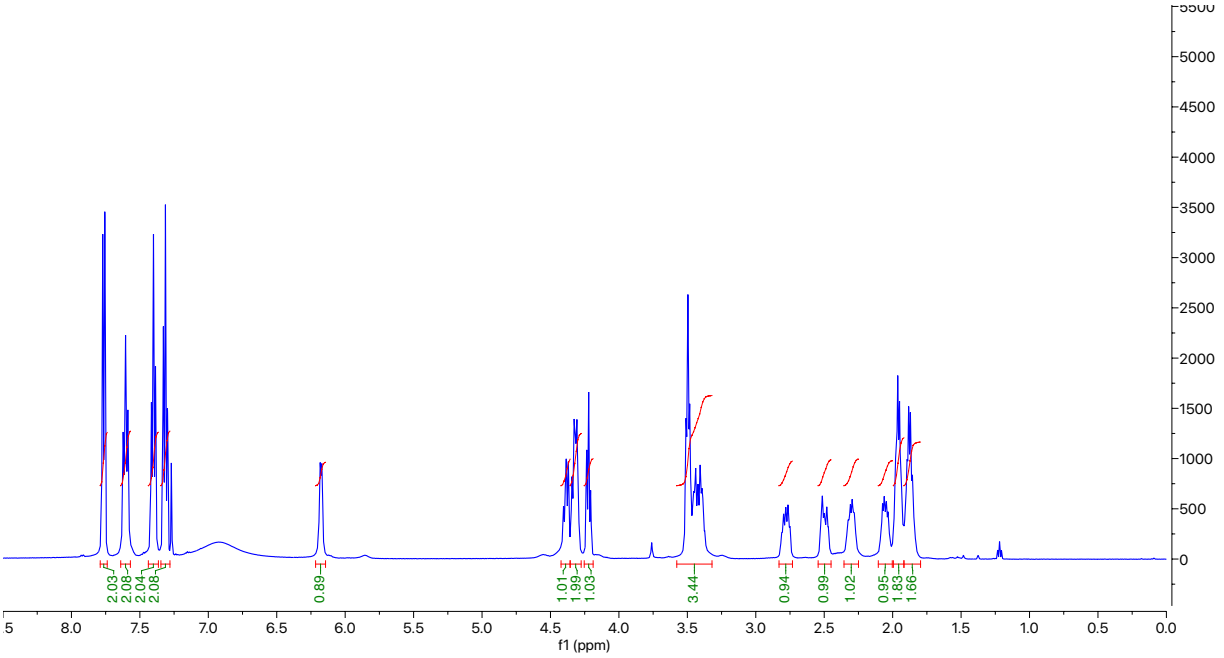


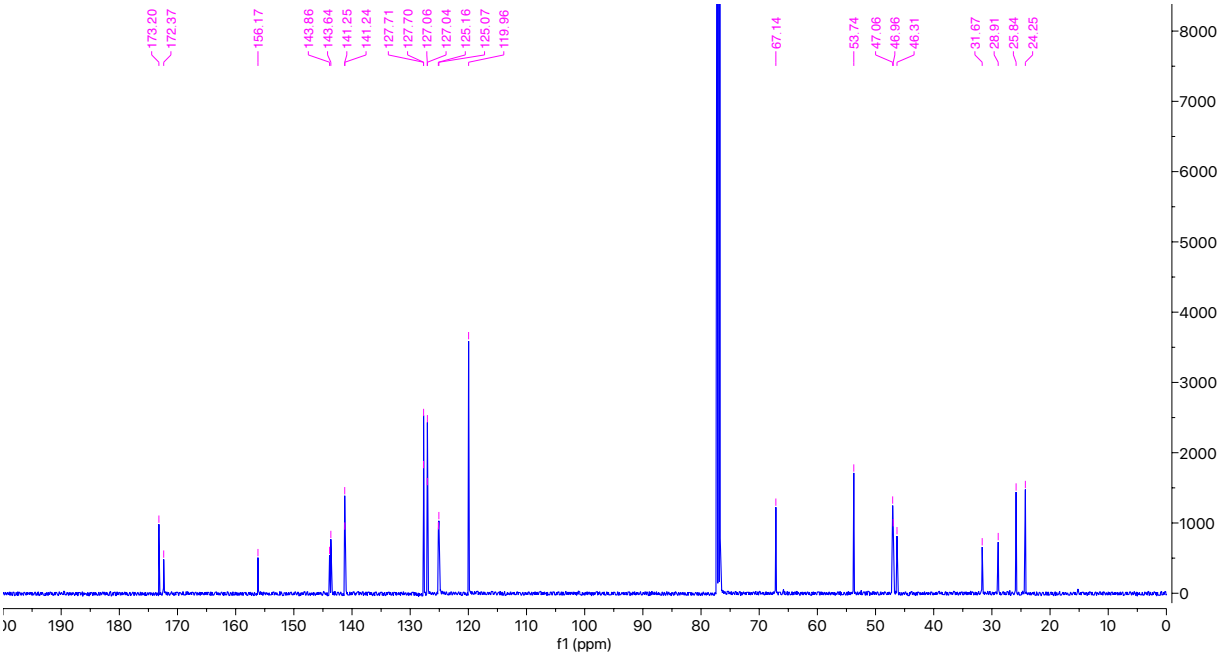


(*S*)-2-((((9*H*-Fluoren-9-yl)methoxy)carbonyl)amino)-5-oxo-5-(pyrrolidin-1-yl)pentanoic acid (L-Fmoc-Pye-OH)

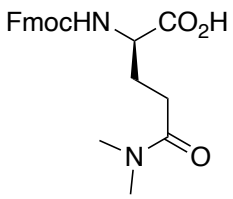


White solid, 88 % yield. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.76 (d, *J* = 7.6 Hz, 2H), 7.60 (t, *J* = 8.6 Hz, 2H), 7.40 (t, *J* = 7.5 Hz, 2H), 7.31 (t, *J* = 7.5 Hz, 2H), 6.18 (d, *J* = 6.2 Hz, 1H), 4.42-4.36 (m, 1H), 4.36-4.28 (m, 2H), 4.22 (t, *J* = 7.2 Hz, 1H), 3.58 – 3.32 (m, 4H), 2.78 (m, 1H), 2.50 (m, 1H), 2.30 (m, 1H), 2.05 (m, 1H), 1.99-1.91 (m, 2H), 1.88 (m, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 173.20, 172.37, 156.17, 143.86, 143.64, 141.25, 141.24, 127.71, 127.70, 127.06, 127.04, 125.16, 125.07, 119.96, 67.14, 53.74, 47.06, 46.96, 46.31, 31.67, 28.91, 25.84, 24.25. HRMS (ESI) m/z calcd for $C_{24}H_{27}N_2O_5^+$ ($M+H^+$) 423.1914; found 423.1913

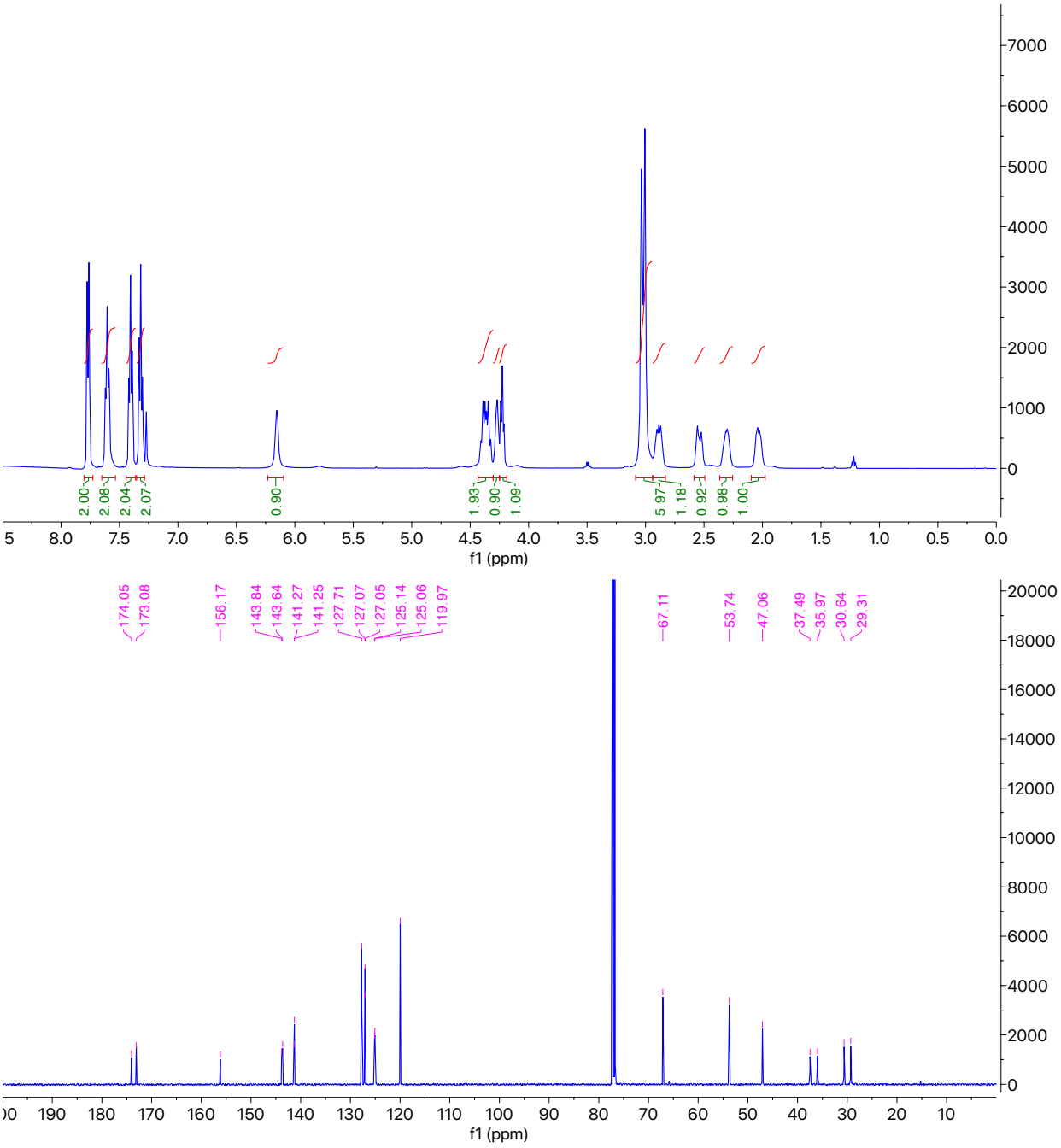




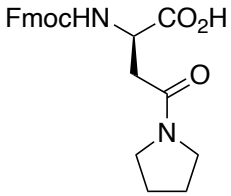
N^2-(((9*H*-Fluoren-9-yl)methoxy)carbonyl)-*N*^{5},*N*^{5}-dimethyl-*D*-glutamine {D-Fmoc-Gln(NMe$_2$)-OH}



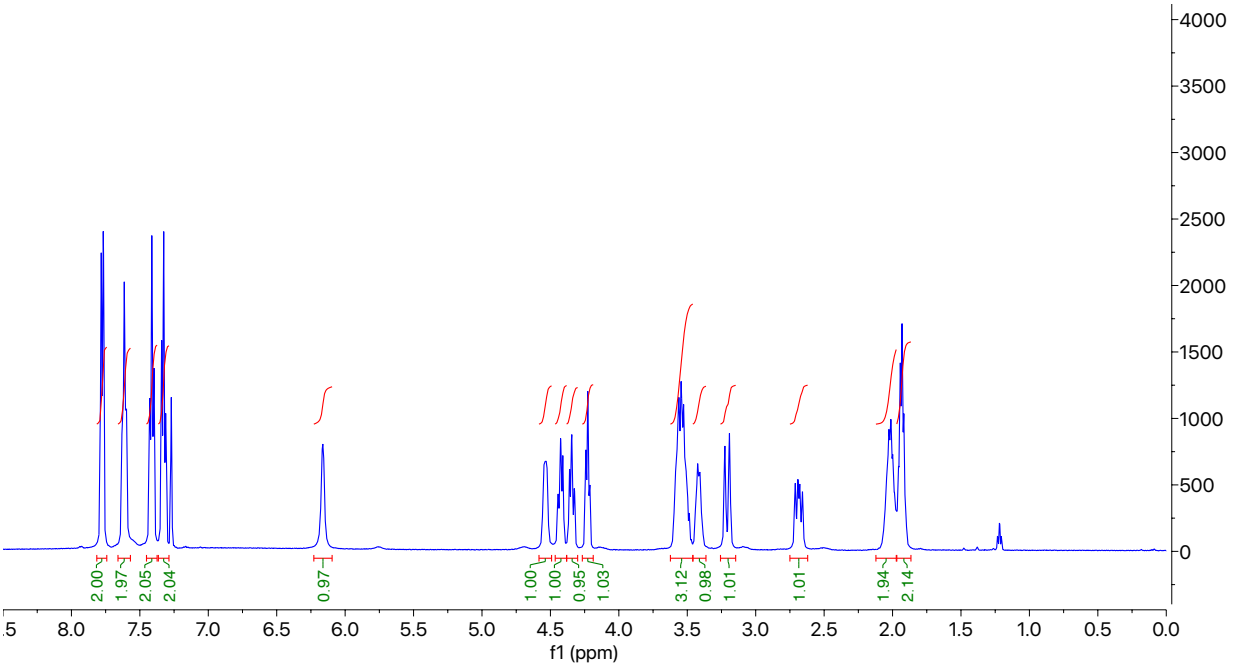
white solid, 85 % yield. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.77 (d, *J* = 7.3 Hz, 2H), 7.61 (t, *J* = 8.0 Hz, 2H), 7.41 (t, *J* = 7.6 Hz, 2H), 7.32 (t, *J* = 7.7 Hz, 2H), 6.15 (br, 1H), 4.43-4.30 (m, 2H), 4.30-4.25 (m, 1H), 4.22 (t, *J* = 7.4 Hz, 1H), 3.03 (s, 3H), 3.01, (s, 3H), 2.89 (m, 1H), 2.55 (m, 1H), 2.31 (m, 1H), 2.03 (m, 1H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 174.05, 173.08, 156.17, 143.84, 143.64, 141.27, 141.25, 127.71, 127.07, 127.05, 125.14, 125.06, 119.97, 67.11, 53.74, 47.06, 37.49, 35.97, 30.64, 29.31. HRMS (ESI) m/z calcd for $C_{22}H_{25}N_2O_5^+$ ($M+H^+$) 397.1758; found 397.1756

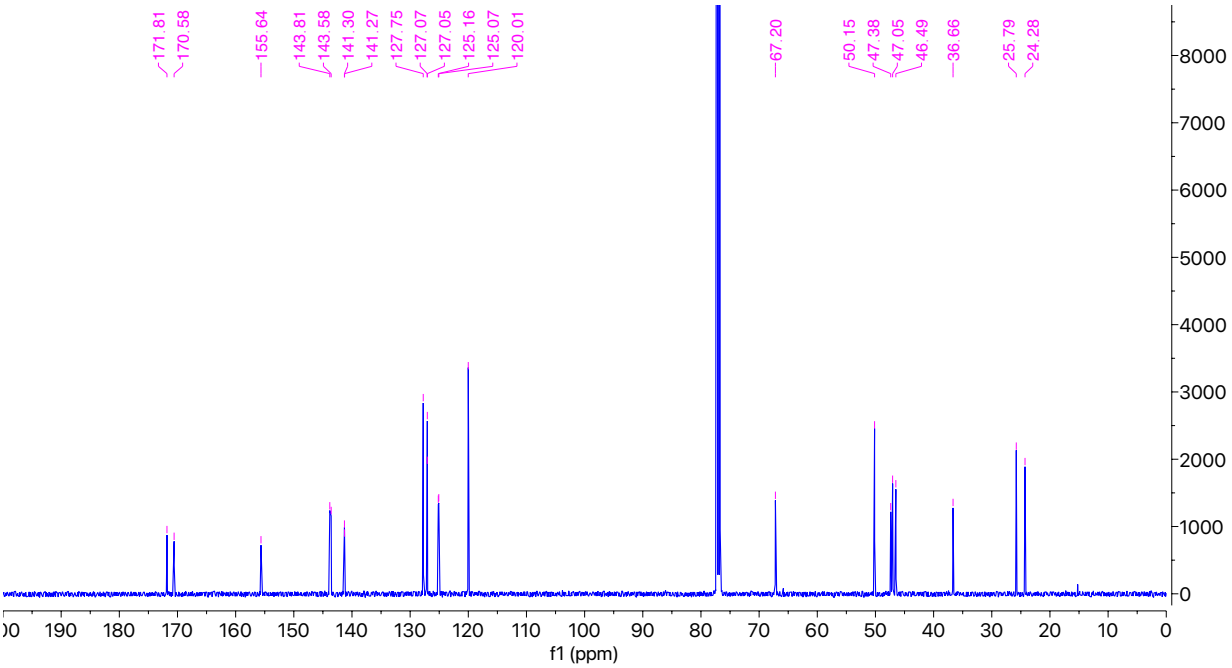


(*R*)-2-((((9*H*-Fluoren-9-yl)methoxy)carbonyl)amino)-4-oxo-4-(pyrrolidin-1-yl)butanoic acid {D-Fmoc-Asn(Pyr)-OH}

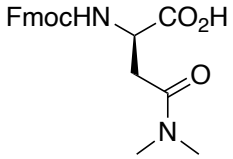


White solid, 84 % yield. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.77 (d, *J* = 7.6 Hz, 2H), 7.61 (t, *J* = 7.4 Hz, 2H), 7.41 (t, *J* = 7.5 Hz, 2H), 7.33 (t, *J* = 7.5, 2.1 Hz, 2H), 6.16 (br, 1H), 4.58-4.49 (m, 1H), 4.42 (t, *J* = 8.2 Hz, 1H), 4.34 (t, *J* = 8.6 Hz, 1H), 4.23 (t, *J* = 7.5 Hz, 1H), 3.62-3.45 (m, 3H), 3.42 (m, 1H), 3.21 (d, *J* = 16.7 Hz, 1H), 2.68 (dd, *J* = 17.0, 9.6 Hz, 1H), 2.11-1.97 (m, 2H), 1.97-1.87 (m, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 171.81, 170.58, 155.64, 143.81, 143.58, 141.30, 141.27, 127.75, 127.07, 127.05, 125.16, 125.07, 120.01, 67.20, 50.15, 47.38, 47.05, 46.49, 36.66, 25.79, 24.28. HRMS (ESI) m/z calcd for $C_{23}H_{25}N_2O_5^+$ ($M+H^+$) 409.1758; found 409.1755

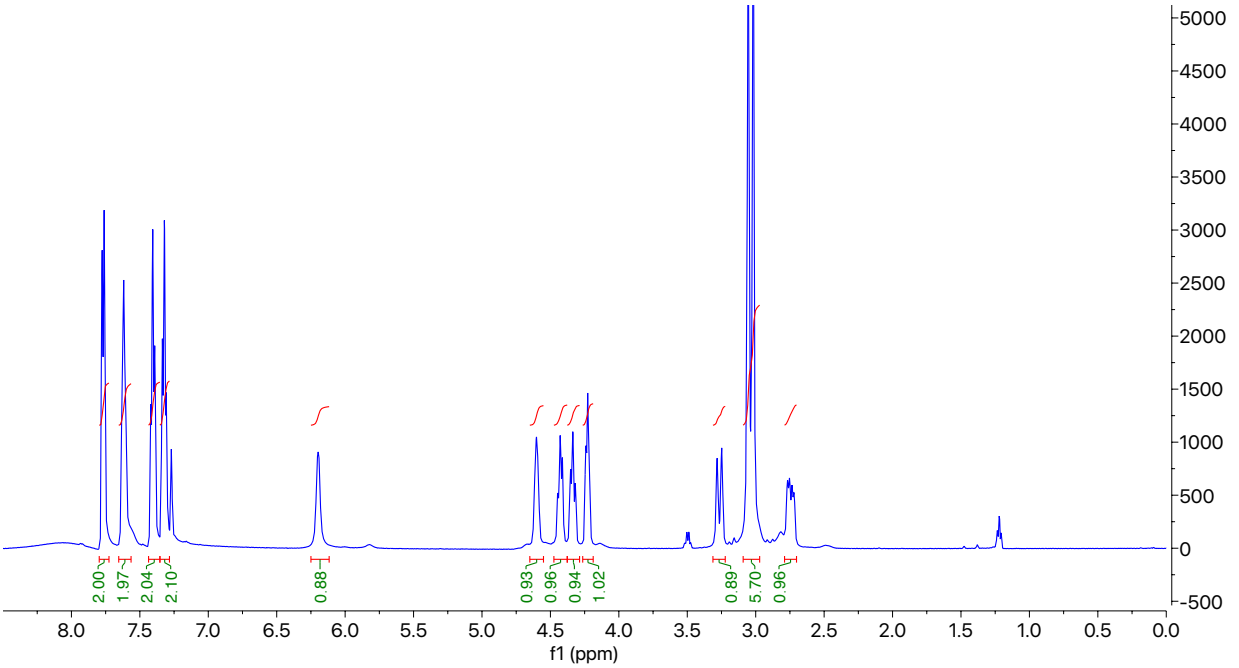


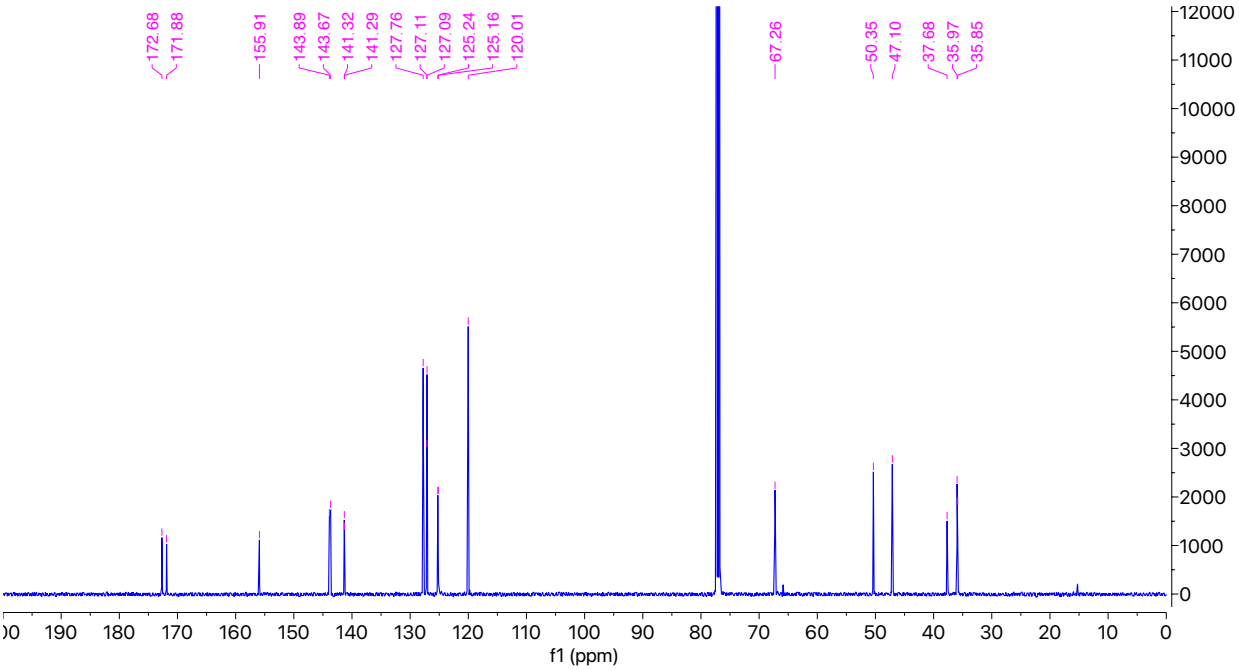


N^2-(((9*H*-Fluoren-9-yl)methoxy)carbonyl)-*N*^{4},*N*^{4}-dimethyl-*D*-asparagine {D-Fmoc-Asn(NMe$_2$)-OH}

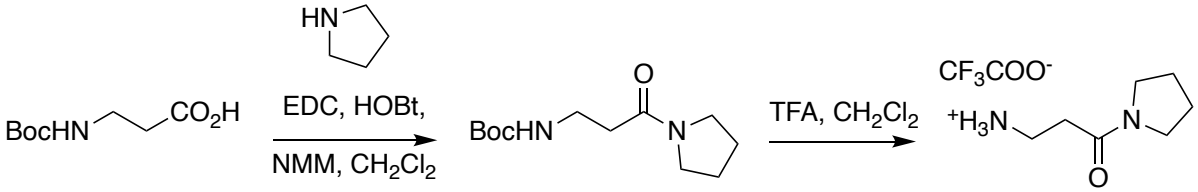


White solid, 79 % yield. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.77 (d, *J* = 7.2 Hz, 2H), 7.61 (t, *J* = 7.3 Hz, 2H), 7.40 (t, *J* = 7.7 Hz, 2H), 7.32 (t, *J* = 7.6 Hz, 2H), 6.20 (br, 1H), 4.65-4.55 (m, 1H), 4.42 (t, *J* = 8.3 Hz, 1H), 4.34 (t, *J* = 9.4 Hz, 1H), 4.26-4.19 (m, 1H), 3.27 (d, *J* = 16.7 Hz, 1H), 3.05 (s, 3H), 3.02 (s, 3H), 2.74 (dd, *J* = 17.2, 7.7 Hz, 1H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 172.68, 171.88, 155.91, 143.89, 143.67, 141.32, 141.29, 127.76, 127.11, 127.09, 125.24, 125.16, 120.01, 67.26, 50.35, 47.10, 37.68, 35.97, 35.85. HRMS (ESI) m/z calcd for $C_{21}H_{23}N_2O_5^+$ (M+H$^+$) 383.1601; found 383.1600



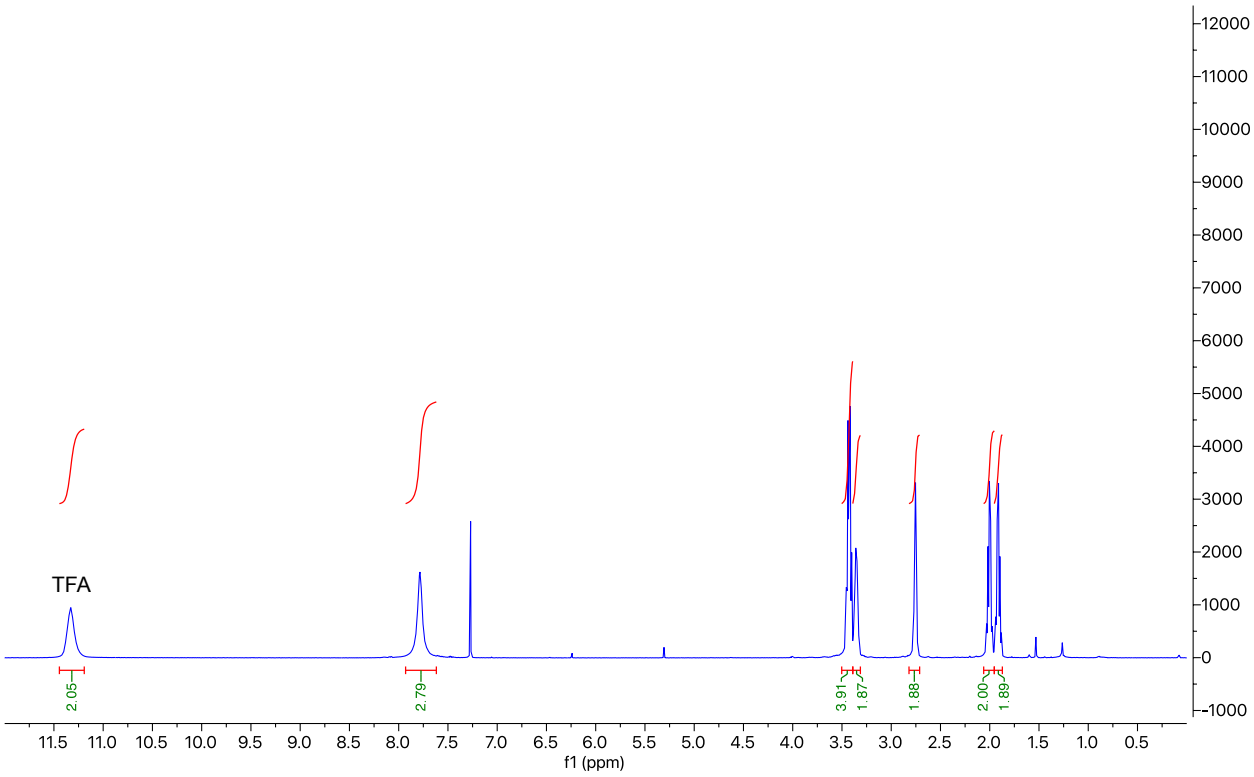


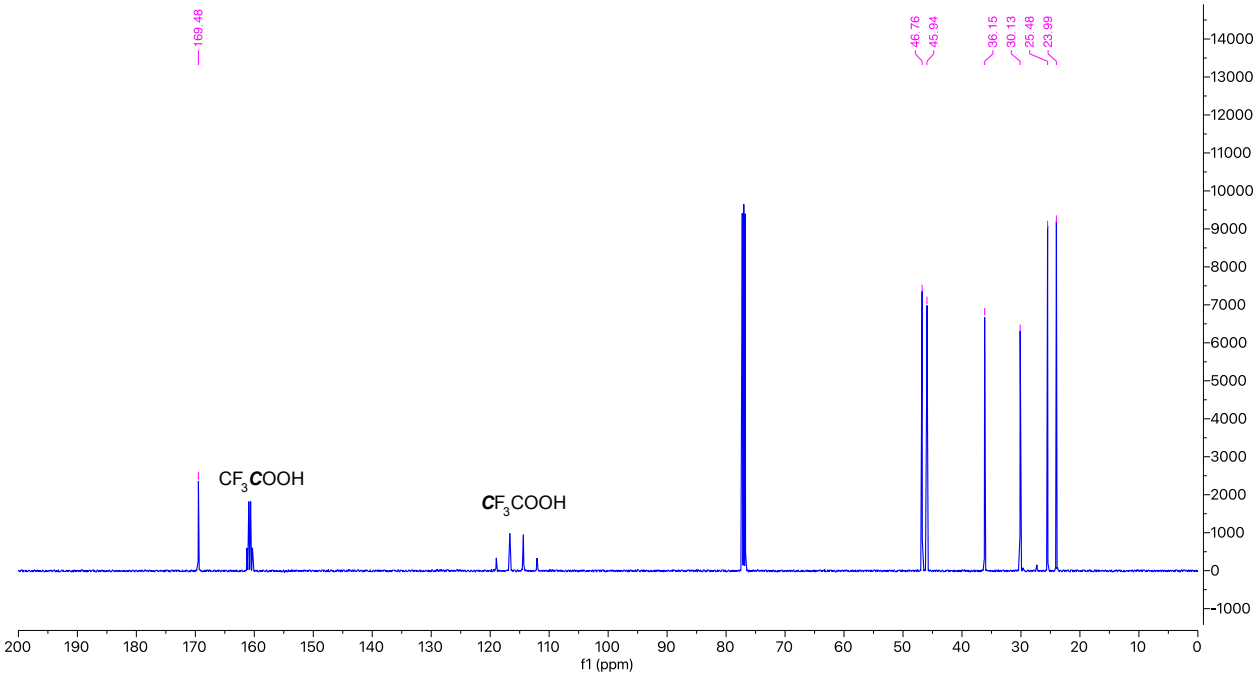
Synthesis of 3-Amino-1-(pyrrolidin-1-yl)propan-1-one



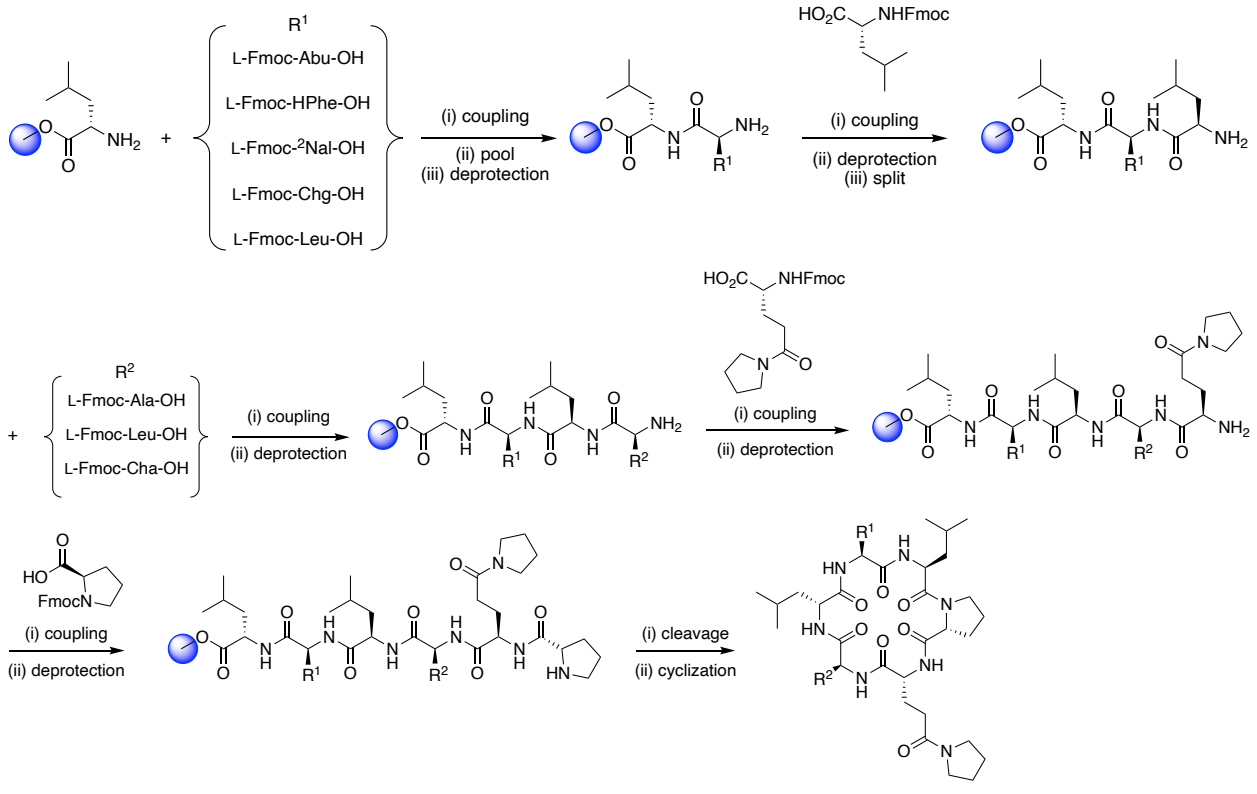
3-((tert-Butoxycarbonyl)amino)propanoic acid (3 mmol), HOBt (4.5 mmol) were dissolved in 30 mL CH_2Cl_2 at 0 ºC. EDC (4.5 mmol) was added and the reaction was stirred for 30 min at 0 ºC. NMM (9 mmol) and pyrrolidine (3.6 mmol) were added subsequently, and the reaction was stirred at ambient temperature overnight. Solution was diluted with CH_2Cl_2 and extracted with 10% citric acid (x 3), Sat. $NaHCO_3$ (x 3) and brine. The organic layer was dried over $MgSO_4$ and reduced in vacuo to obtain the product. This crude material was stirred in 1:1 TFA/CH_2Cl_2 solution (20 mL) for 3 h. After removing excess solvent, product was dried under vacuum and used for the synthesis of peptoid analogue.

yellow oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ c7.78 (s, 3H), 3.50-3.39 (m, 4H), 3.39-3.32 (m, 2H), 2.75 (t, *J* = 5.6 Hz, 2H), 2.00 (quint, *J* = 6.8 Hz, 2H), 1.91 (quint, *J* = 6.8 Hz, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 169.48, 46.76, 45.94, 36.15, 30.13, 25.48, 23.99.

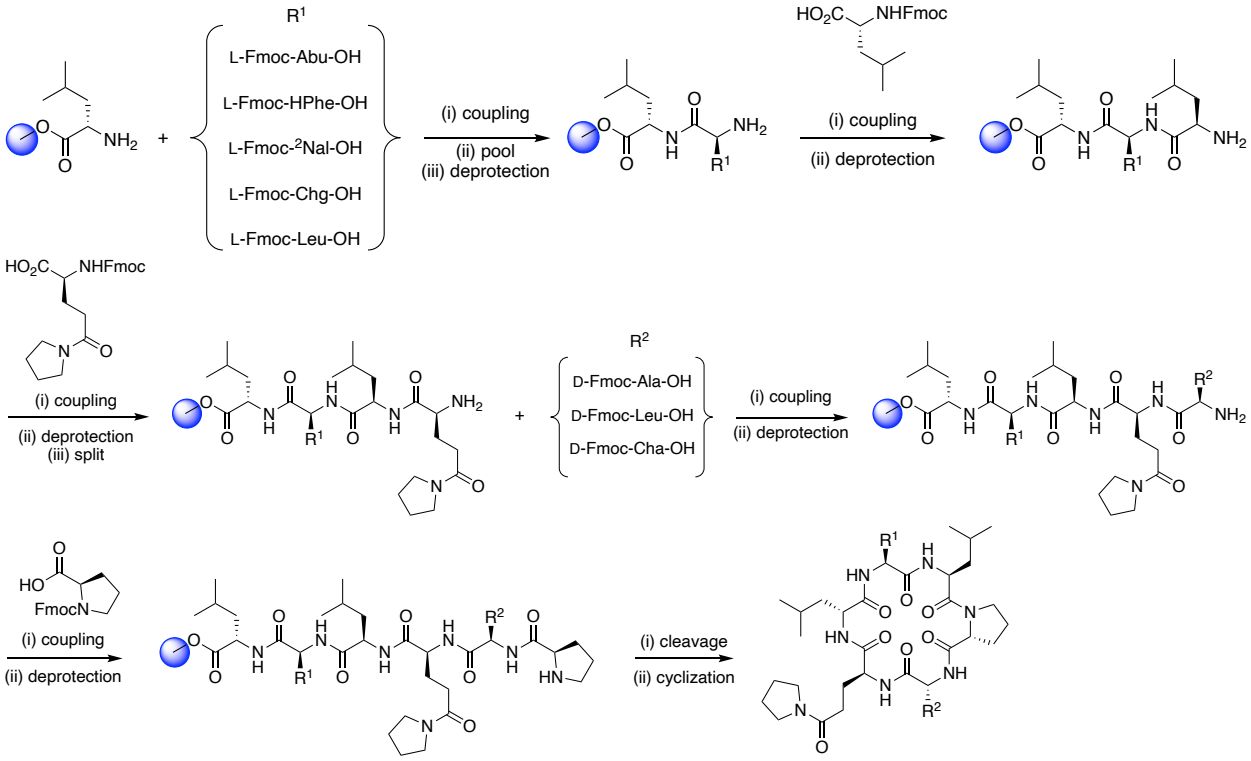




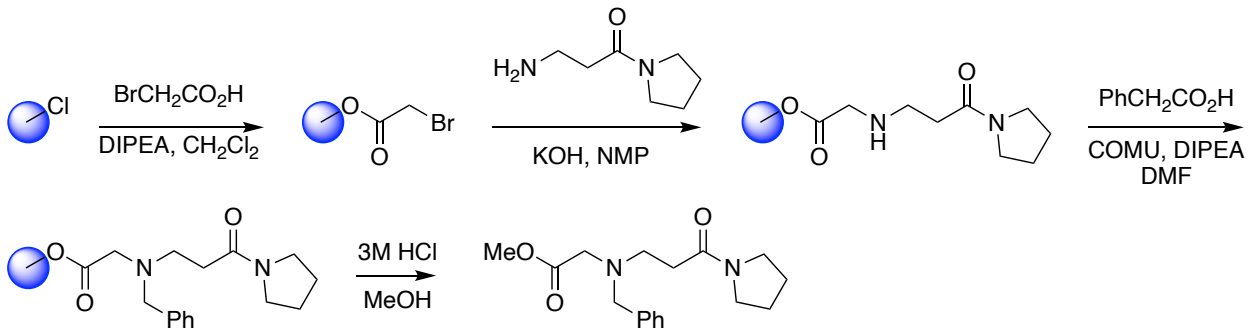
Synthesis Scheme of Cyclic Peptides and Peptoid Monomer



Scheme S3 Synthesis scheme of **3-Pye²** library



Scheme S4 Synthesis scheme of **3-Pye**[3] library



Scheme S5 Synthesis scheme of Pye peptoid monomer.

In Silico Investigation of Side Chain-to-Backbone Hydrogen Bonding of Model Structures

Distance between side chain HBA and the amide HBD of model structures was investigated *in silico* using Discovery Studio as described. SMILE strings of each designed model were obtained from Chemdraw, then 3D structures were generated by Discovery Studio software. A conformer algorithm based on the Merck Molecular Force Field (MMFF) was used to generate the conformations of each model. Initial structures were minimized, then conformation ensembles were generated with SMART minimization method available as the default setting in Discovery Studio. Implicit solvation was incorporated using the Generalized Born model with the dielectric constant set to 4.81 to represent chloroform solvent. Duplicate structures within RMSD < 0.2 Å were discarded. The distance between the *O* acceptor from the side chain and the N*H* acetyl donor was measured and plotted against conformer's energy relative to the lowest energy conformer. Molecules with short HBD-HBA distance at low relative energy conformation have the potential to form intramolecular hydrogen bonding in a low dielectric environment.

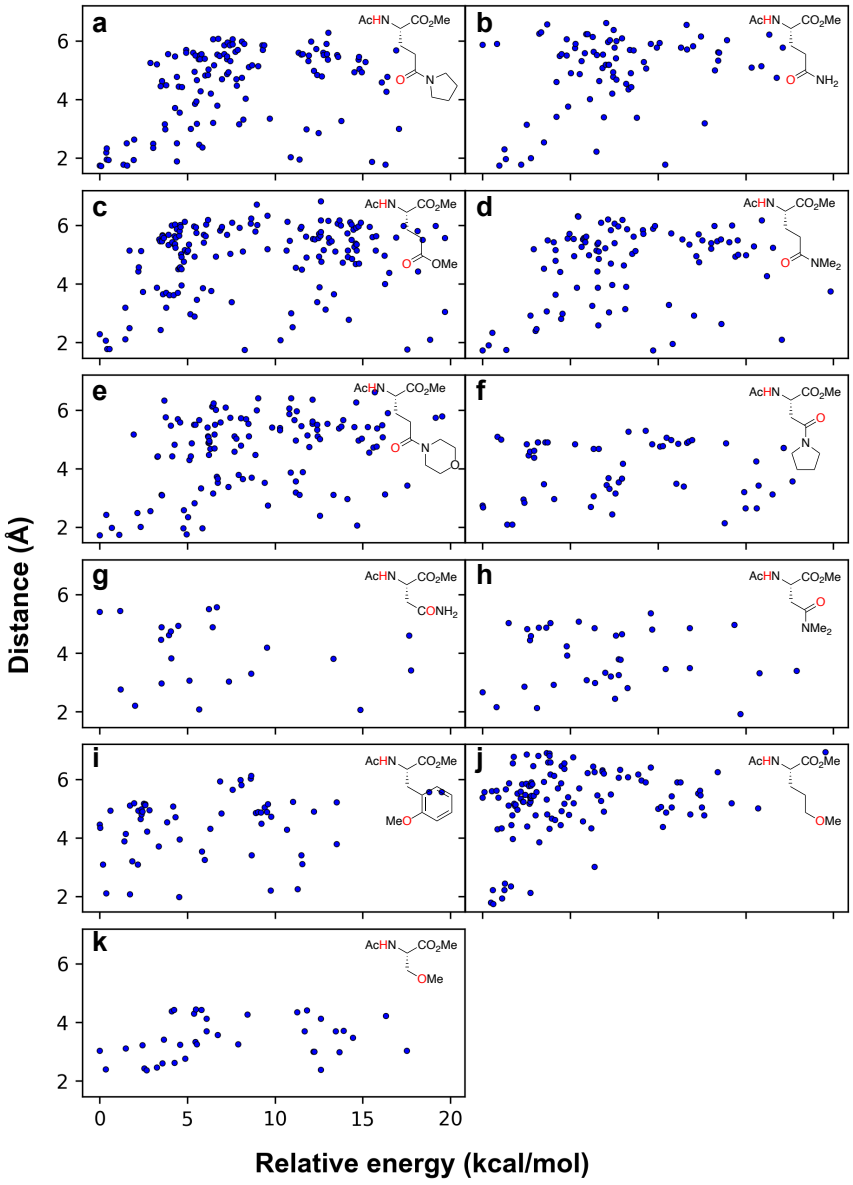


Figure S1. Plots between HBD-HBA distances of each conformer and their relative energy.

Analytical Procedures

General Protocol of UPLC-MS Analysis

UPLC-MS analyses of cyclic peptides were performed via a Thermo Scientific Ultimate 3000 UPLC system, using a Thermo Hypersil GOLD C18 30 mm x 2.1 mm (1.9 µm) column (#25002-032130). Flow rate was set to 1 mL/min and a gradient method was as followed: 0.0-0.5 min, 5% MeCN; 0.5-0.75 min, ramp to 95% MeCN; 0.75-3.0 min, 95% MeCN; 3.0-3.5 min, 5% MeCN. Mass identification and quantification used an inline Thermo Scientific Orbitrap VelosPro (FTMS mode), tuned for maximum ionization of cyclosporin A, background ion locking on octyl phthalate, 200-2000 AMU mass windows, using +/- 0.02 AMU windows for integration.

Shake Flask Partition Experiment

Shake flask partition experiment to measure log $D_{dec/w}$ was performed as described by Naylor, M. R. *et*. *al*.[1] with minor modification. 1,9-Decadiene and PBS buffer (pH 7.4) were saturated by vortexing with an equal volume for 1-2 min, then centrifuged to allow the emulsion to separate. 8 µL of 400 µM test compounds in DMSO solution were added to 2.0 mL Eppendorf tube followed by 800 µL of saturated 1,9-decadiene and 800 µL saturated PBS buffer to yield a final concentration of 2 µM with 0.5% DMSO. Tubes were sealed carefully and agitated by vortex (30 min) and sonication (30 min). Tubes were then centrifuged for 10 min at 16,000 x g to separate two layers. 150 µL x 4 of each layer was carefully separated to minimize contamination from the pipette tip, especially when collecting the lower aqueous layer: 1) expel an air while passing through organic layer to prevent capillary wicking of organic solvent, 2) gently wipe off organic solvent adhered to tip after collecting and before dispensing. Solutions were collected in 96-well 300 µL conical-bottom plate as quadruplicate for each layer and evaporated overnight in a Genevac centrifugal evaporator (60 °C). Each sample was resuspended with 150 uL of 1:1 MeCN/H_2O, mixed, sealed, and sonicated for 30 min. After a final centrifugation for 10 min at 700 x g, samples were quantified via UPLC-MS in 10uL injections with gradient as described above. The average integrations from quadruplicate data of each layer were used to calculate the partition coefficient (log $D_{dec/w}$) of 1,9-decadiene/water. Carbamazepine was used as a standard in these experiments (log $D_{dec/w}$ = -0.18). LPE was calculated from the equation published previously.[1]

$$\text{LPE} = \log D_{dec/w} - 1.06\text{ALogP} + 5.47$$

Whereas ALogP is a 2D molecular descriptor to represent the octanol/water partition coefficient determined from atoms in a molecule, which was calculated from Discovery Studio software.

Parallel Artificial Membrane Permeability Assay (PAMPA)

A 96-well donor plate with 0.45 µm hydrophobic Immobilon-P membrane supports (Millipore MAIPNTR10) and a 96-well Teflon acceptor plate (Millipore MSSACCEPTOR) were used in the PAMPA permeability test. The acceptor plate was prepared by adding 300 µL of 5% DMSO in PBS buffer (pH 7.4) to each well. Sample solutions were prepared by diluting DMSO stock solutions to a final 2 µM sample concentration in PBS buffer with a final DMSO concentration of 5%.

A 1% (w/v) solution of lecithin (soybean, 90%) in *n*-dodecane was prepared and sonicated before use. 5 µL of the *n*-dodecane / lecithin solution was carefully applied to the underside of membrane supports in the wells of the donor plate, with care taken to not touch the pipet tip to the membrane. After approximately 15 minutes, 150 µL of the 2 µM test compounds were added to the donor wells. The donor plate was then placed on top of the acceptor plate so that the artificial membrane was in contact with the buffer solution below, ensuring that no bubbles form beneath the membrane. A lid was placed on the donor well, and the whole assembly was covered within a sealed chamber and left overnight at room temperature. A wet paper towel was placed inside the chamber to prevent evaporation. After ~15 h (exact time recorded and used for subsequent calculations) the donor and acceptor plates were separated, and 50 µL of each well (donor and acceptor) were mixed with 50 µL MeOH in another 96-well 300 µL conical-bottom plate and sealed. These solutions were analyzed as quadruplicate via UPLC-MS as described above. Permeability (%T) was quantified as the ratio of

analyte area in the acceptor well divided by a theoretical equilibrium ratio based on amounts of combined analyte found in the donor and acceptor wells as follows:

$$\boldsymbol{AnalyteEquil} = \frac{(Ia * Va) + (Id * Vd)}{Va + Vd}$$

$$\%\boldsymbol{T} = \left(\frac{Ia}{[\boldsymbol{AnalyteEquil}]}\right) * 100$$

Recovery (%R) was quantified as the ratio of total compound identified in the donor and acceptor wells relative to the total compound identified in the original dilution sample.

$$\%\boldsymbol{R} = \frac{(Ia * Va) + (Id * Vd)}{Ir * Vd} * 100$$

Permeation rates (P_{app}) were calculated from %T by the following equations:

$$\boldsymbol{C} = \frac{Vd * Va}{(Vd + Va) * Msa * Ts}$$

$$\boldsymbol{P_{app}}\ (cm/s) = -\boldsymbol{C} * \ln\left(1 - \left(\frac{\%\boldsymbol{T}}{100}\right)\right)$$

Where:

- Active surface area of membrane (mm^2): **Msa** = 240
- Volume of acceptor well (µL): **Va** = 300
- Volume of donor well (µL): **Vd** = 150
- Assay run time (s): **Ts**
- Donor intensity: **Id**
- Acceptor intensity: **Ia**
- Recovery intensity: **Ir**

In addition, propranolol was used as a standard reference to confirm the assay was performed correctly, which had log P_{app} ~ -5.

Thermodynamic Solubility Assay

20 µL of 10 mM stock solutions were dispensed into a 96-well conical plate and evaporated overnight in a Genevac centrifugal evaporator (60 °C). 100 uL of PBS (pH 7.4) was reintroduced to the plate to yield a maximum 2mM concentration, then the plate was sealed and sonicated for 1 h. The plate was then gently agitated at 37 °C for ~24 h. The mixtures were filtered through a 0.7 µm glass fiber filter plate (Agilent 200965-100) into a 96-well conical plate. The filtrate can be further diluted with MeCN in a new 96-well plate (dilution factor can be up to 40) and sealed for quantification via UPLC-MS. Standard curves of each compound were acquired from serial dilution of stock solution with DMSO (50 µM to 0.1 µM) and used to calculate concentrations of analytes. All standards and analytes were performed in triplicate and averaged.

Table S1. Experimental Data of Pye-scanning Cyclic Hexapeptides

ID	log $D_{dec/w}$	LPE	PAMPA P_{app} x 10^{-6} (cm/s)	PAMPA log P_{app}	Solubility (μM)	Caco-2 log P_{app} (cm/s)
1	2.03	3.54	8.11 ± 0.28	-5.09 ± 0.02	16 ± 3	
1-Pye^2	1.26	4.12	1.72 ± 0.08	-5.77 ± 0.02	640 ± 4	
1-Pye^3	-0.41	2.45	0.66 ± 0.07	-6.19 ± 0.05	545 ± 97	
1-Pye^4	0.02	2.89	0.30 ± 0.01	-6.53 ± 0.01	580 ± 54	
1-Pye^5	-0.42	2.45	0.32 ± 0.03	-6.50 ± 0.05	505 ± 21	
1-Pye^6	-0.30	2.56	0.66 ± 0.19	-6.20 ± 0.14	720 ± 27	
2	1.96	3.47	5.26 ± 1.85	-5.30 ± 0.14	241 ± 20	
2-Pye^2	-0.29	2.58	0.56 ± 0.11	-6.26 ± 0.08	607 ± 22	
2-Pye^3	-0.29	2.58	1.09 ± 0.46	-5.99 ± 0.17	446 ± 16	
2-Pye^4	0.04	2.91	1.29 ± 0.89	-5.96 ± 0.27	456 ± 91	
2-Pye^5	-0.82	2.04	0.13 ± 0.03	-6.89 ± 0.09	553 ± 44	
2-Pye^6	-0.29	2.58	1.50 ± 0.86	-5.89 ± 0.28	562 ± 92	
3	2.04	3.55	7.98 ± 0.41	-5.10 ± 0.02	63 ± 6	-5.58
3-Pye^2	1.06	3.93	1.46 ± 0.09	-5.84 ± 0.03	734 ± 29	-6.41
3-Pye^3	-0.22	2.65	0.23 ± 0.02	-6.64 ± 0.03	144 ± 20	-7.15
3-Pye^4	0.21	3.07	0.14 ± 0.01	-6.86 ± 0.04	578 ± 125	-6.77
3-Pye^5	0.14	3.00	0.28 ± 0.02	-6.55 ± 0.03	554 ± 17	-6.70
3-Pye^6	0.16	3.02	0.18 ± 0.02	-6.74 ± 0.06	598 ± 11	

Table S2. Experimental Data of 3-Pye^2 Library

ID	MW	R^1	R^2	RT (sec)	ALogP	log $D_{dec/w}$	P_{app} x 10^{-6} cm/s	log P_{app}	% recovery	solubility (μM)
3-Pye^2(Ala^{3}Abu^5)	661.42	Abu	Ala	97.3	0.516	-1.4	0.030 ± 0.002	-7.53 ± 0.02	96.3 ± 3.6	
3-Pye^2(Ala^3)	689.45	Leu	Ala	100.1	1.224	-0.010	0.21 ± 0.02	-6.68 ± 0.05	89.5 ± 3.1	
3-Pye^2(Abu^5)	703.46	Abu	Leu	100.5	1.748	0.060	0.20 ± 0.01	-6.71 ± 0.02	94.6 ± 2.0	
3-Pye^2(Ala^{3}Chg^5)	715.46	Chg	Ala	101.9	1.832	0.46	0.78 ± 0.05	-6.11 ± 0.03	84.7 ± 0.6	
3-Pye^2(Ala^{3}HPhe^5)	737.45	HPhe	Ala	101.9	2.004	0.48	1.25 ± 0.10	-5.90 ± 0.03	87.9 ± 0.4	
3-Pye^2	731.50	Leu	Leu	103.6	2.456	1.2	1.28 ± 0.04	-5.89 ± 0.01	63.1 ± 2.5	
3-Pye^2(Ala^{3}Nal^5)	773.45	^{2}Nal	Ala	103.0	2.457	0.90	3.04 ± 0.33	-5.52 ± 0.05	52.5 ± 2.6	
3-Pye^2(Cha^{3}Abu^5)	743.50	Abu	Cha	103.5	2.745	1.1	2.05 ± 0.10	-5.69 ± 0.02	87.3 ± 6.9	
3-Pye^2(Chg^5)	757.51	Chg	Leu	105.7	3.064	1.6	5.94 ± 0.62	-5.23 ± 0.05	88.1 ± 4.4	
3-Pye^2(HPhe^5)	779.50	HPhe	Leu	105.3	3.236	1.5	5.65 ± 0.54	-5.25 ± 0.04	104.1 ± 4.9	82 ± 6
3-Pye^2(Cha^3)	771.53	Leu	Cha	106.6	3.453	2.1	8.93 ± 0.68	-5.05 ± 0.03	70.6 ± 5.5	111 ± 29
3-Pye^2(Nal^5)	815.50	^{2}Nal	Leu	106.4	3.688	2.1	8.21 ± 1.73	-5.09 ± 0.10	58.3 ± 6.7	67 ± 44
3-Pye^2(Cha^{3}Chg^5)	797.54	Chg	Cha	109.5	4.061	2.6	14.11 ± 1.20	-4.85 ± 0.04	59.2 ± 3.9	74 ± 10
3-Pye^2(Cha^{3}HPhe^5)	819.53	HPhe	Cha	108.4	4.233	2.5	13.57 ± 1.09	-4.87 ± 0.03	69.9 ± 3.9	4.0 ± 0.4
3-Pye^2(Cha^{3}Nal^5)	855.53	^{2}Nal	Cha	109.9	4.685	3.0	5.36 ± 1.08	-5.28 ± 0.09	24.2 ± 1.3	N.D.

N.D. = not detectable

Table S3. Experimental Data of 3-Pye^3 Library

ID	MW	R^1	R^2	RT (sec)	ALogP	log $D_{dec/w}$	P_{app} x 10^{-6} cm/s	log P_{app}	% recovery
3-Pye^3(Ala^{2}Abu^5)	661.42	Abu	Ala	98.0	0.516	-1.9	0.004 ± 0.001	-8.39 ± 0.12	98.9 ± 2.9
3-Pye^3(Ala^2)	689.45	Leu	Ala	100.8	1.224	-1.2	0.030 ± 0.003	-7.50 ± 0.05	81.0 ± 3.5

3-Pye^3(Abu^5)	703.46	Abu	Leu	100.7	1.748	-0.25	0.03 ± 0.01	-7.46 ± 0.08	71.9 ± 3.5
3-Pye^3(Ala^{2}Chg^5)	715.46	Chg	Ala	102.9	1.832	-0.73	0.15 ± 0.03	-6.83 ± 0.09	86.5 ± 2.8
3-Pye^3(Ala^{2}HPhe^5)	737.45	HPhe	Ala	102.2	2.004	-0.57	0.25 ± 0.06	-6.61 ± 0.10	89.6 ± 2.8
3-Pye^3	731.50	Leu	Leu	103.2	2.456	0.18	0.20 ± 0.02	-6.70 ± 0.03	50.0 ± 0.6
3-Pye^3(Ala^{2}Nal^5)	773.45	^{2}Nal	Ala	102.9	2.457	-0.090	0.52 ± 0.11	-6.29 ± 0.09	70.5 ± 2.9
3-Pye^3(Cha^{2}Abu^5)	743.50	Abu	Cha	104.3	2.745	0.34	0.41 ± 0.03	-6.39 ± 0.03	60.2 ± 4.4
3-Pye^3(Chg^5)	757.51	Chg	Leu	105.3	3.064	0.63	0.64 ± 0.10	-6.20 ± 0.07	79.0 ± 3.6
3-Pye^3(HPhe^5)	779.50	HPhe	Leu	104.8	3.236	0.55	1.06 ± 0.17	-5.98 ± 0.07	79.3 ± 2.9
3-Pye^3(Cha^2)	771.53	Leu	Cha	107.6	3.453	1.2	2.22 ± 0.31	-5.66 ± 0.06	46.5 ± 1.2
3-Pye^3(Nal^5)	815.50	^{2}Nal	Leu	104.9	3.688	1.2	1.89 ± 0.39	-5.73 ± 0.09	56.8 ± 3.0
3-Pye^3(Cha^{2}Chg^5)	797.54	Chg	Cha	109.2	4.061	1.7	3.35 ± 0.86	-5.49 ± 0.12	56.4 ± 3.7
3-Pye^3(Cha^{2}HPhe^5)	819.53	HPhe	Cha	108.5	4.233	1.6	4.15 ± 1.19	-5.40 ± 0.14	56.7 ± 5.1
3-Pye^3(Cha^{2}Nal^5)	855.53	^{2}Nal	Cha	108.5	4.685	2.2	4.04 ± 1.67	-5.43 ± 0.22	26.3 ± 3.0

Table S4. Experimental Data of Cyclic Hexapeptides with Different HBAs

ID	ALogP	log $D_{dec/w}$	LPE	PAMPA P_{app} x 10^{-6} (cm/s)	PAMPA log P_{app}	Solubility (μM)
3-Glu(OMe)^2	2.46	0.69	3.6	1.97 ± 0.56	-5.72 ± 0.13	714 ± 28
3-Gln(NMe$_2$)^2	2.00	0.43	3.8	0.70 ± 0.29	-6.18 ± 0.19	1657 ± 29
3-Asn(Pyr)^2	2.13	0.53	3.7	0.56 ± 0.01	-6.25 ± 0.01	669 ± 67
3-Asn(NMe$_2$)^2	1.67	0.24	3.9	0.28 ± 0.02	-6.56 ± 0.03	1118 ± 69

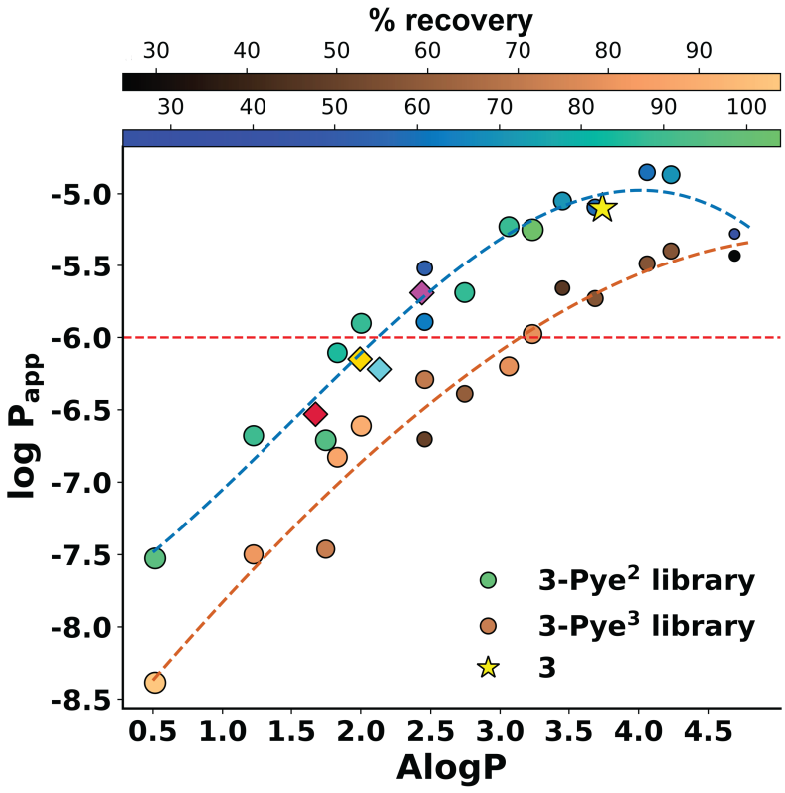
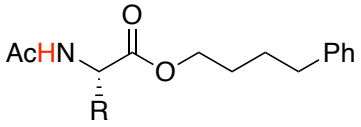
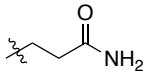
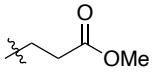
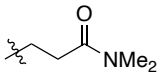
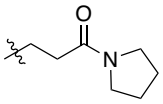
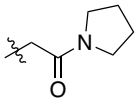
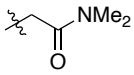


Figure S2. Comparison of PAMPA permeability data of **3-Glu(OMe)^2** (magenta diamond), **3-Asn(Pyr)^2** (blue diamond), **3-Gln(NMe$_2$)^2** (yellow diamond) and **3-Asn(NMe$_2$)^2** (red diamond) to the **3-Pye^2** and **3-Pye^3** libraries.

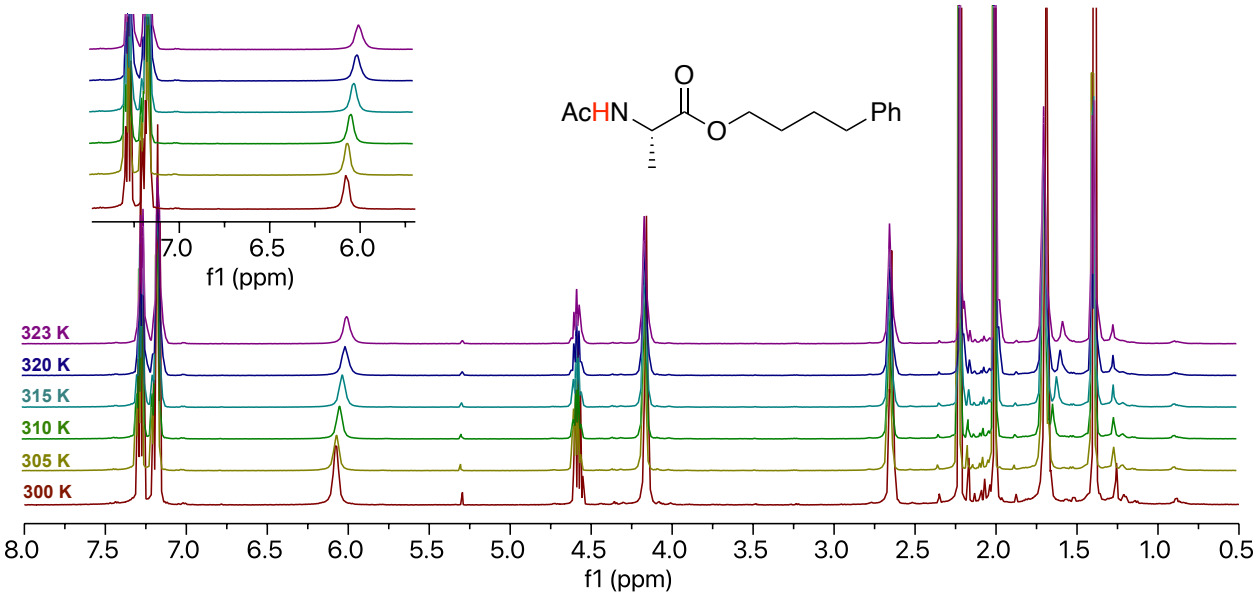
NMR Spectra of Amide Proton Temperature Coefficient Experiment



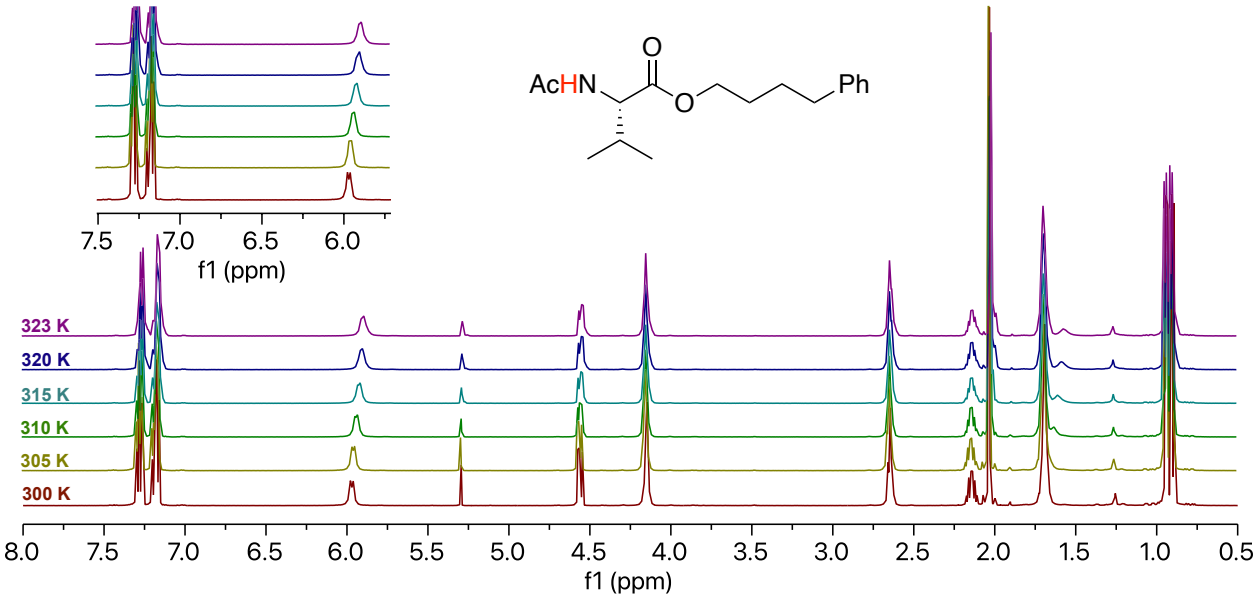
4

ID	R	δ_{NH} (300 K)	$\delta_{NH}/\Delta T$ (ppb/K)
4-Ala	-CH_3	6.08	-3.54
4-Val		5.97	-3.31
4-Leu		5.86	-3.02
4-Ile		5.96	-2.95
4-Gln		6.59	-5.37
4-Glu(OMe)		6.18	-3.51
4-Gln(NMe_2)		6.88	-6.48
4-Gln(Pyr)		7.11	-7.84
4-Asn(Pyr)		6.81	-1.66
4-Asn(NMe_2)		6.89	-1.52

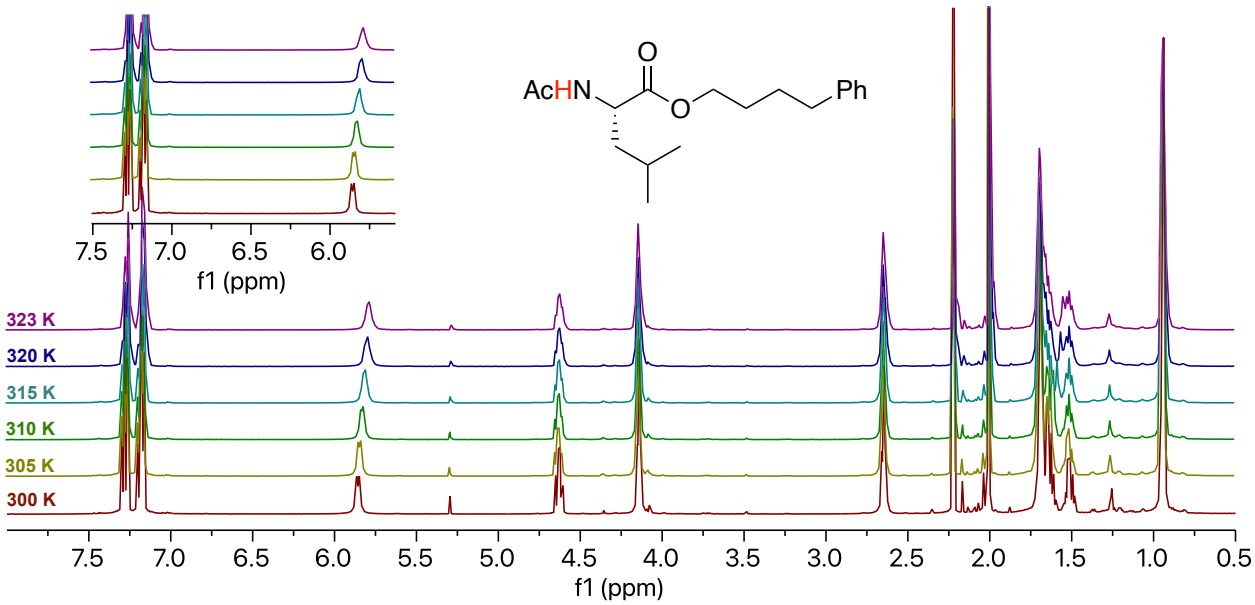
4-Ala


AcHN
O
O
Ph
323 K
320 K
315 K
310 K
305 K
300 K
7.0
6.5
6.0
f1 (ppm)
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)

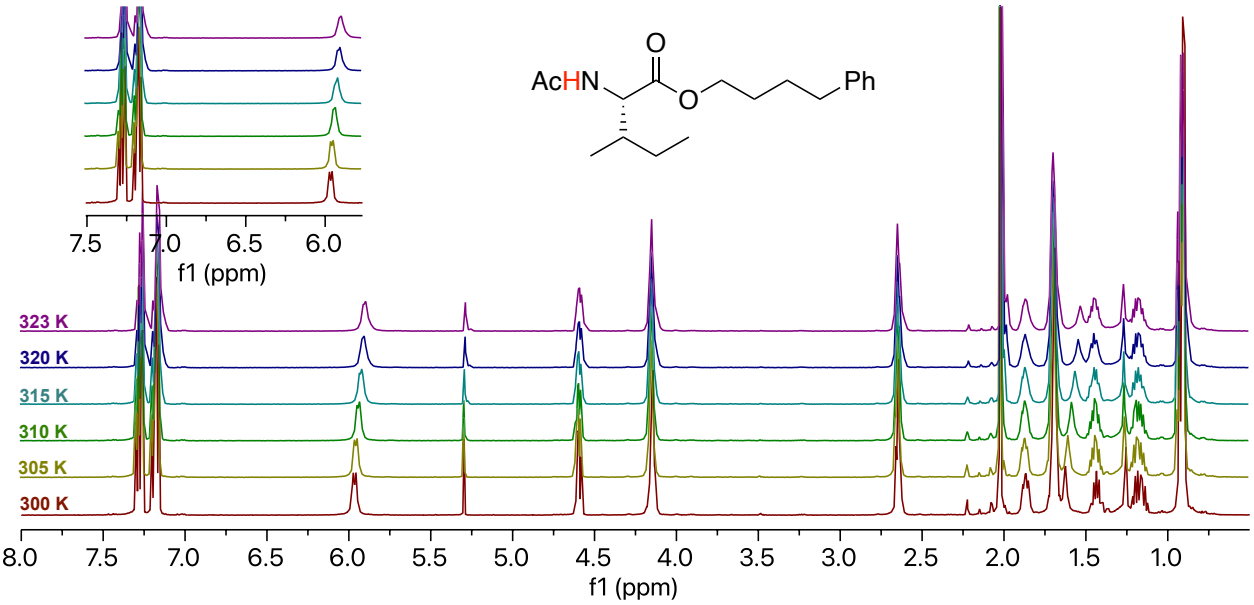
4-Val


AcHN
O
O
Ph
323 K
320 K
315 K
310 K
305 K
300 K
7.5
7.0
6.5
6.0
f1 (ppm)
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)

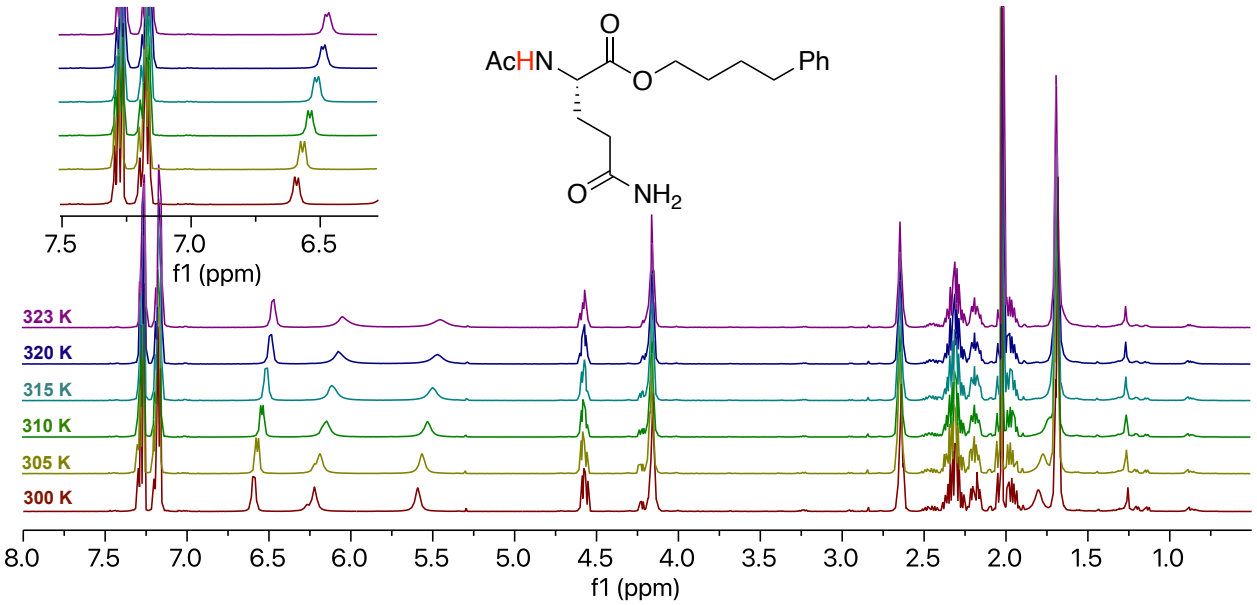
4-Leu


AcHN
O
O
Ph
323 K
320 K
315 K
310 K
305 K
300 K
7.5 7.0 6.5 6.0
f1 (ppm)
7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5
f1 (ppm)

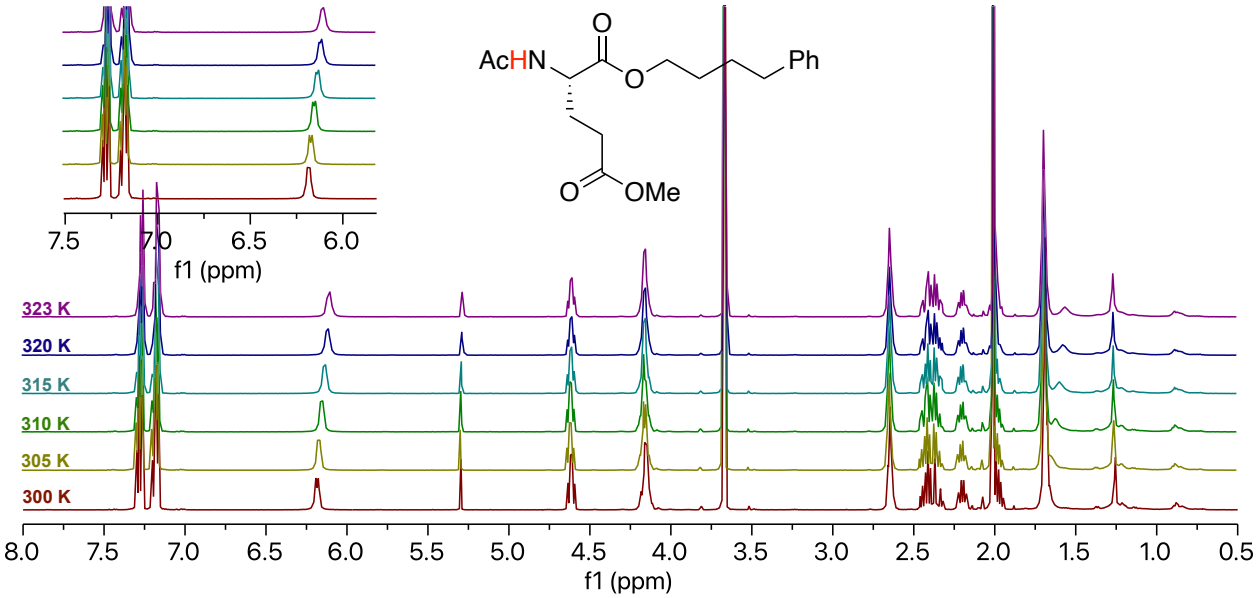
4-Ile


AcHN
O
O
Ph
323 K
320 K
315 K
310 K
305 K
300 K
7.5 7.0 6.5 6.0
f1 (ppm)
8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0
f1 (ppm)

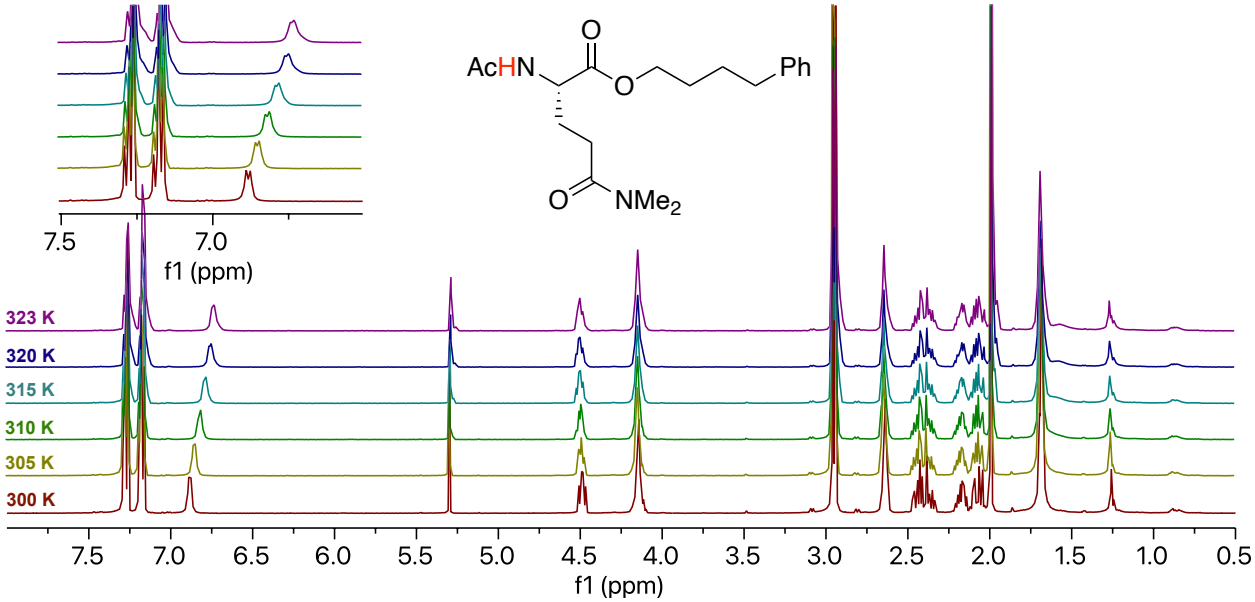
4-Gln



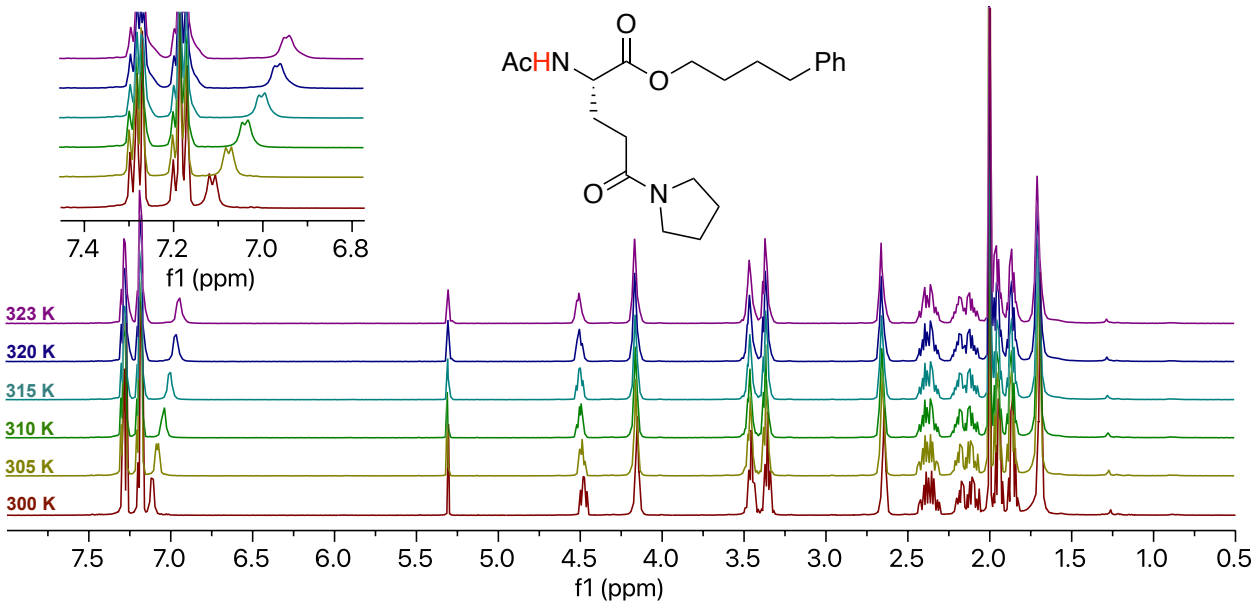
4-Glu(OMe)



4-Gln(NMe$_2$)



4-Gln(Pyr)



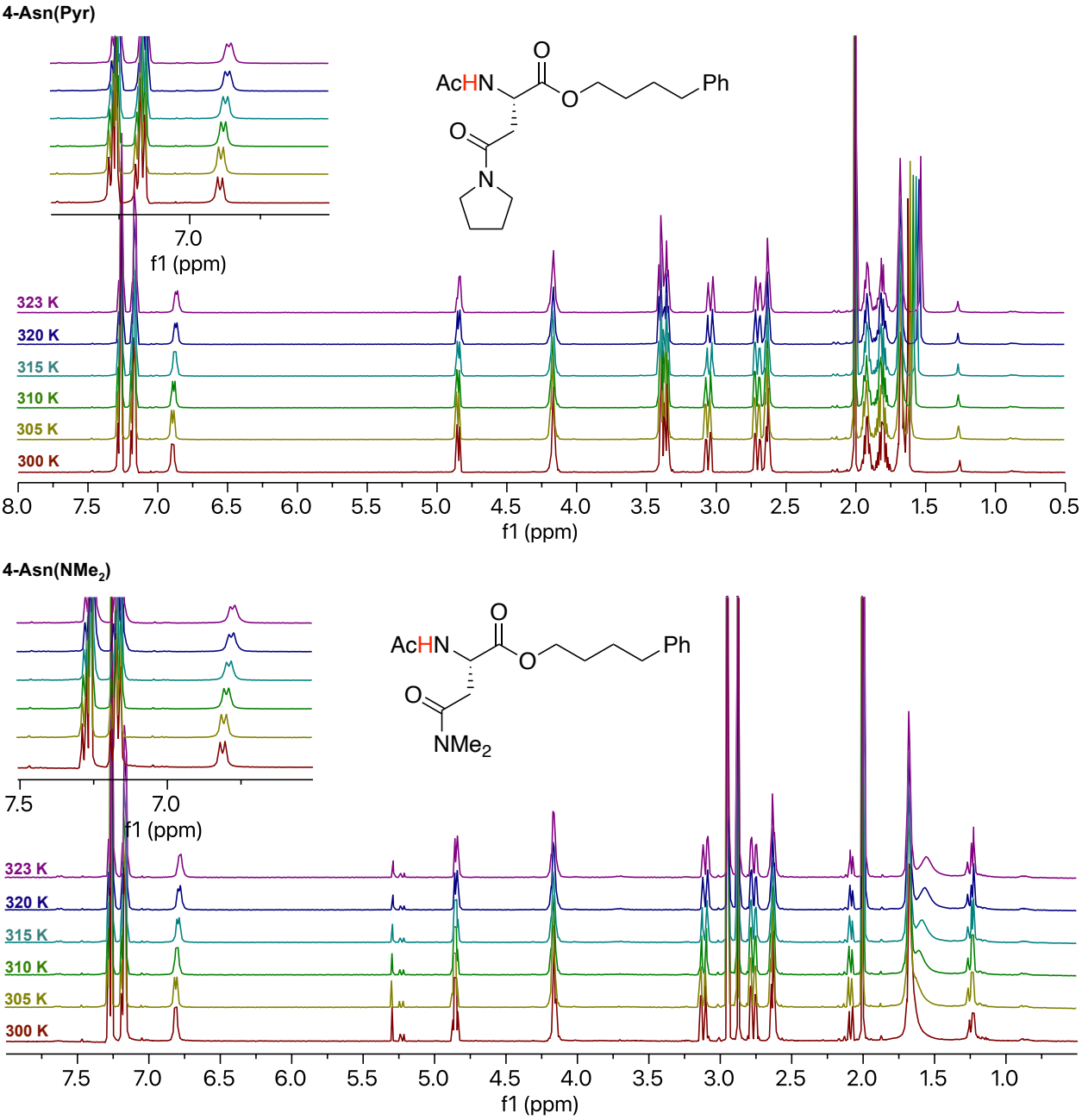


Figure S3. ^{1}H NMR of model structures **4** at different temperatures.

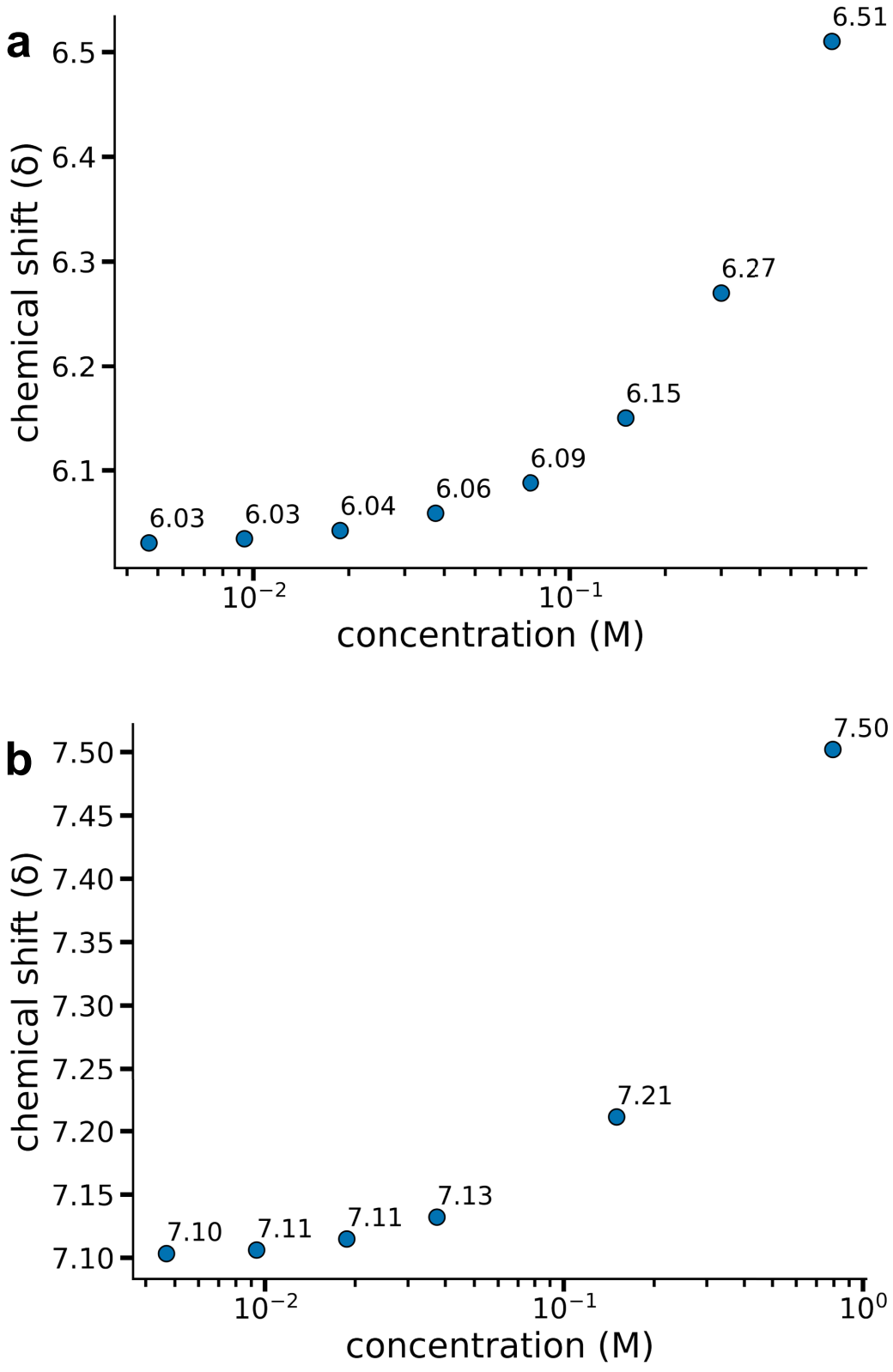
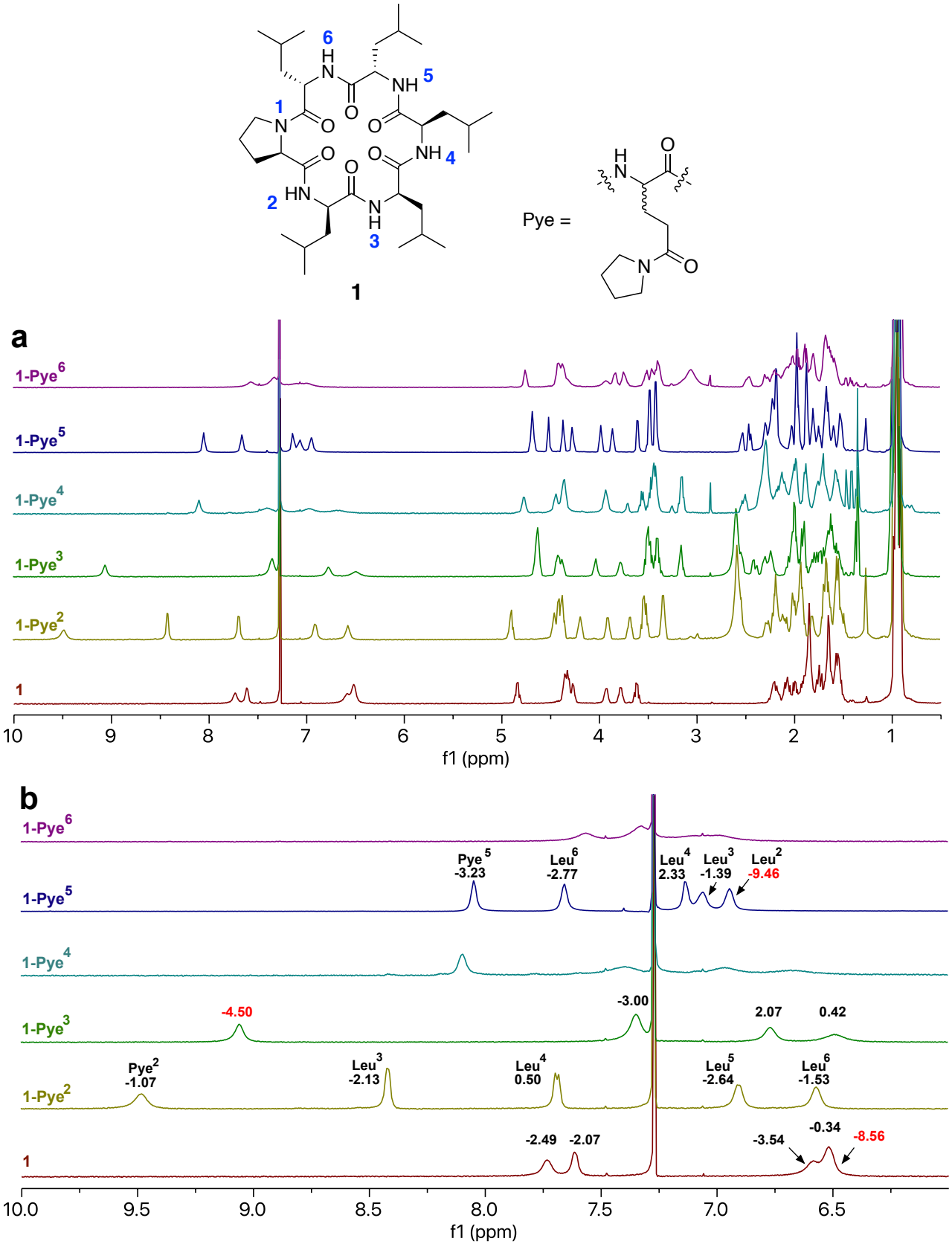
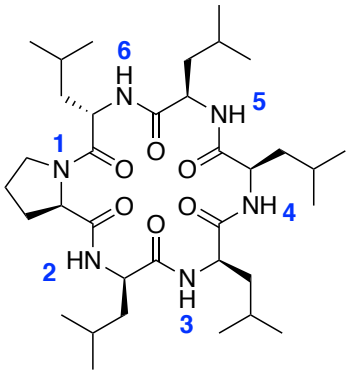


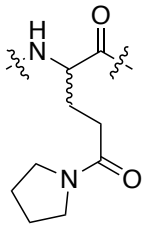
Figure S4. Amide NH chemical shift of (**a**) **4-Ala** and (**b**) **4-Gln(Pyr)** at various concentrations.

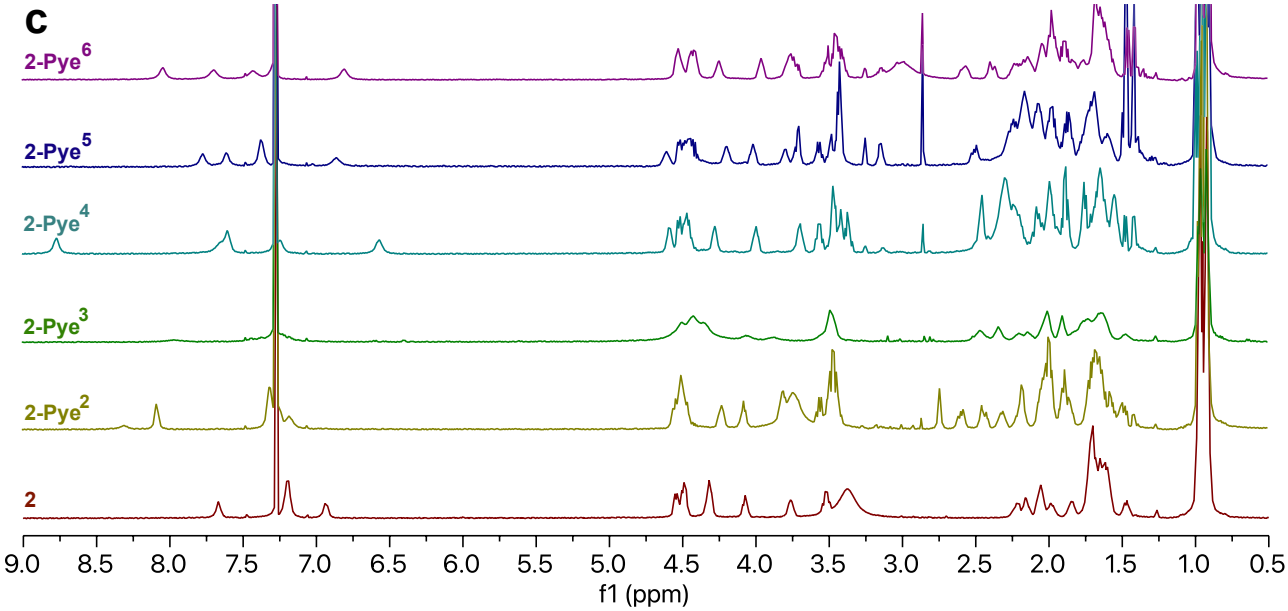

1
Pye =
a
1-Pye6
1-Pye5
1-Pye4
1-Pye3
1-Pye2
1
10
9
8
7
6
5
4
3
2
1
f1 (ppm)
b
1-Pye6
1-Pye5
1-Pye4
1-Pye3
1-Pye2
1
Pye5 -3.23
Leu6 -2.77
Leu4 2.33
Leu3 -1.39
Leu2 -9.46
-4.50
-3.00
2.07
0.42
Pye2 -1.07
Leu3 -2.13
Leu4 0.50
Leu5 -2.64
Leu6 -1.53
-2.49
-2.07
-3.54
-0.34
-8.56
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
f1 (ppm)

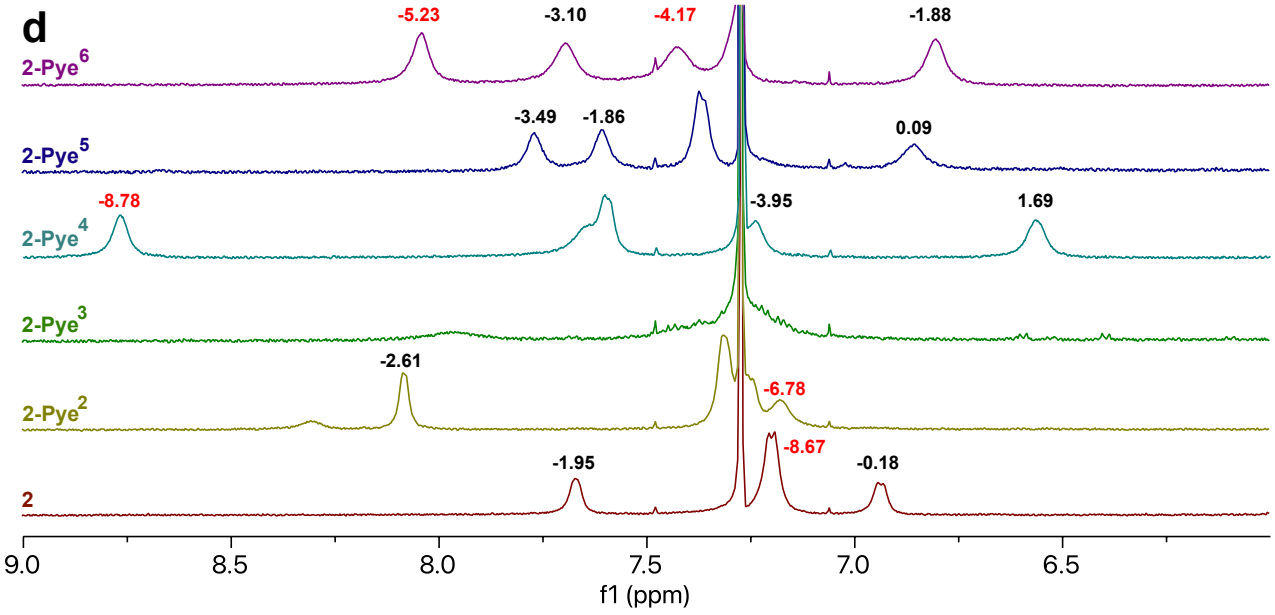

6
H
N
5
NH
1
N
O
O
O
O
O
NH
4
2
HN
O
N
H
3

2

Pye =


H
N
O
N
O


c
2-Pye6
2-Pye5
2-Pye4
2-Pye3
2-Pye2
2
9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5
f1 (ppm)


d
-5.23
-3.10
-4.17
-1.88
2-Pye6
-3.49
-1.86
0.09
2-Pye5
-8.78
-3.95
1.69
2-Pye4
2-Pye3
-2.61
-6.78
-8.67
2-Pye2
-1.95
-0.18
2
9.0 8.5 8.0 7.5 7.0 6.5
f1 (ppm)

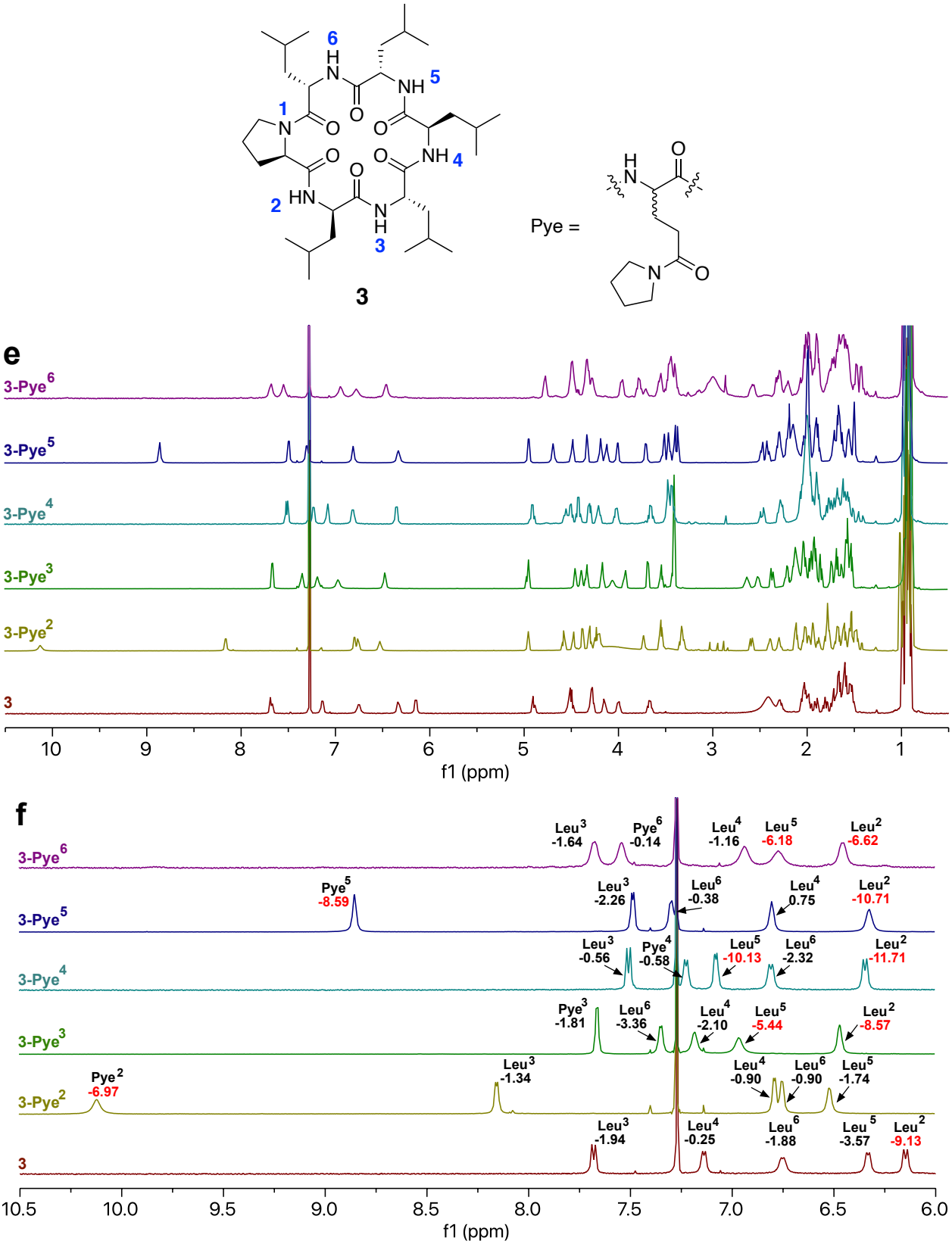
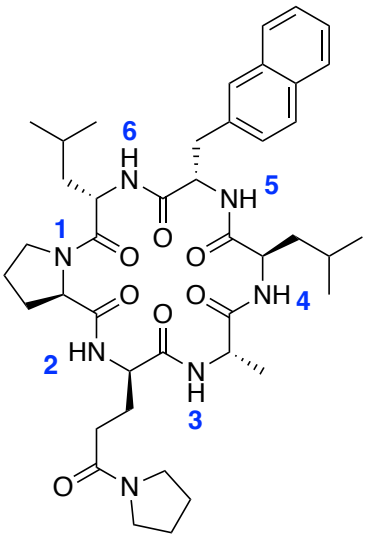
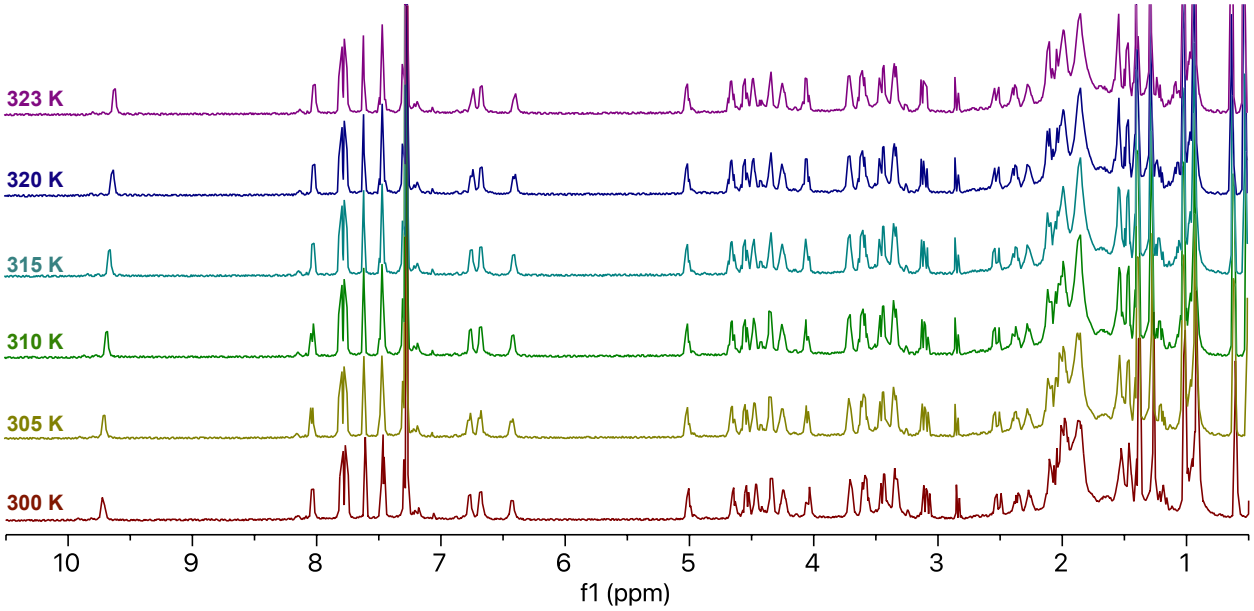


Figure S5 ^{1}H NMR of cyclic hexapeptides **1**, **2**, **3** and their derivatives at ~10 mM. Values indicate the amide temperature coefficients in ppb/K.



3-Pye^{2}(Ala^{3}Nal^{5})



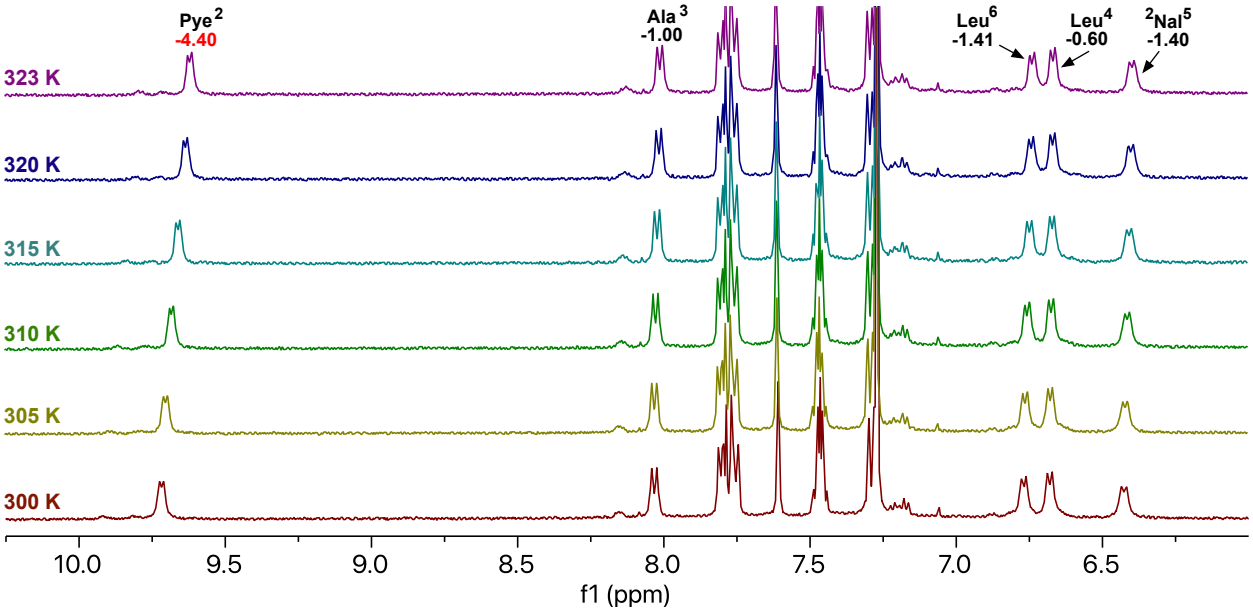
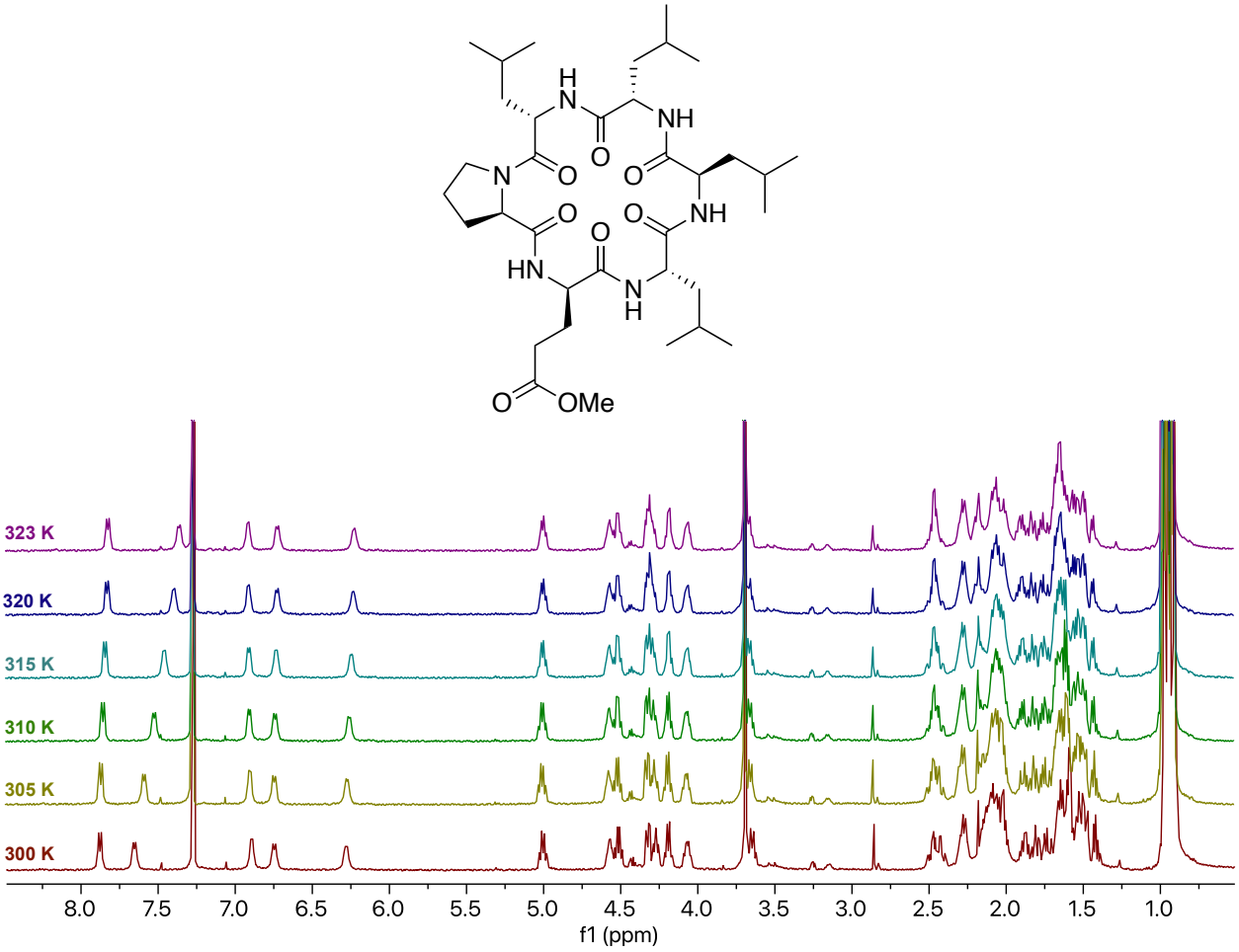


Figure S6. ^{1}H NMR of **3-Pye^{2}(Ala^{3}Nal^{5})**.



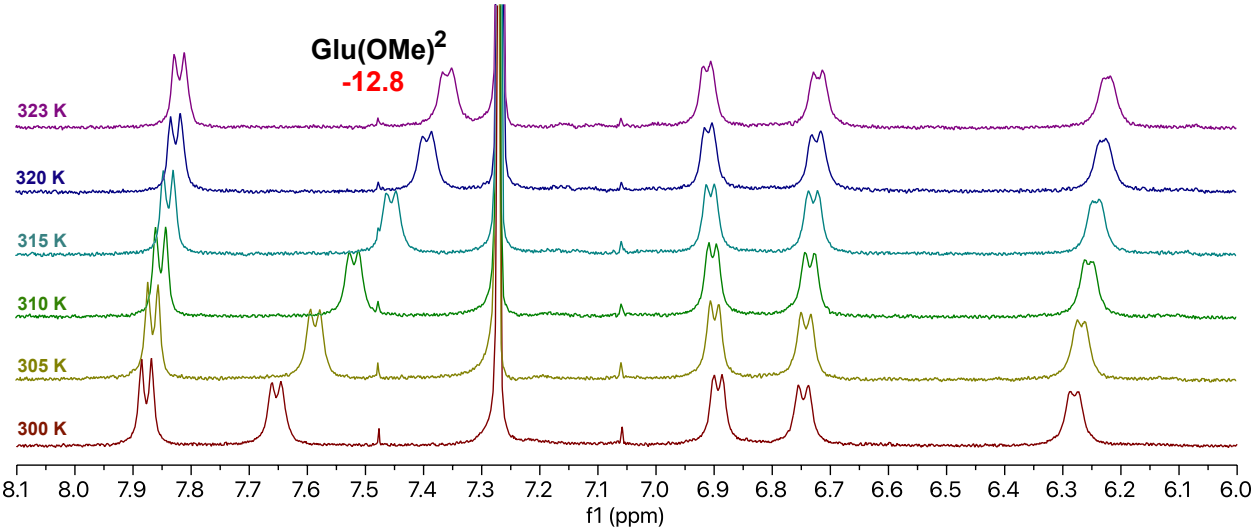
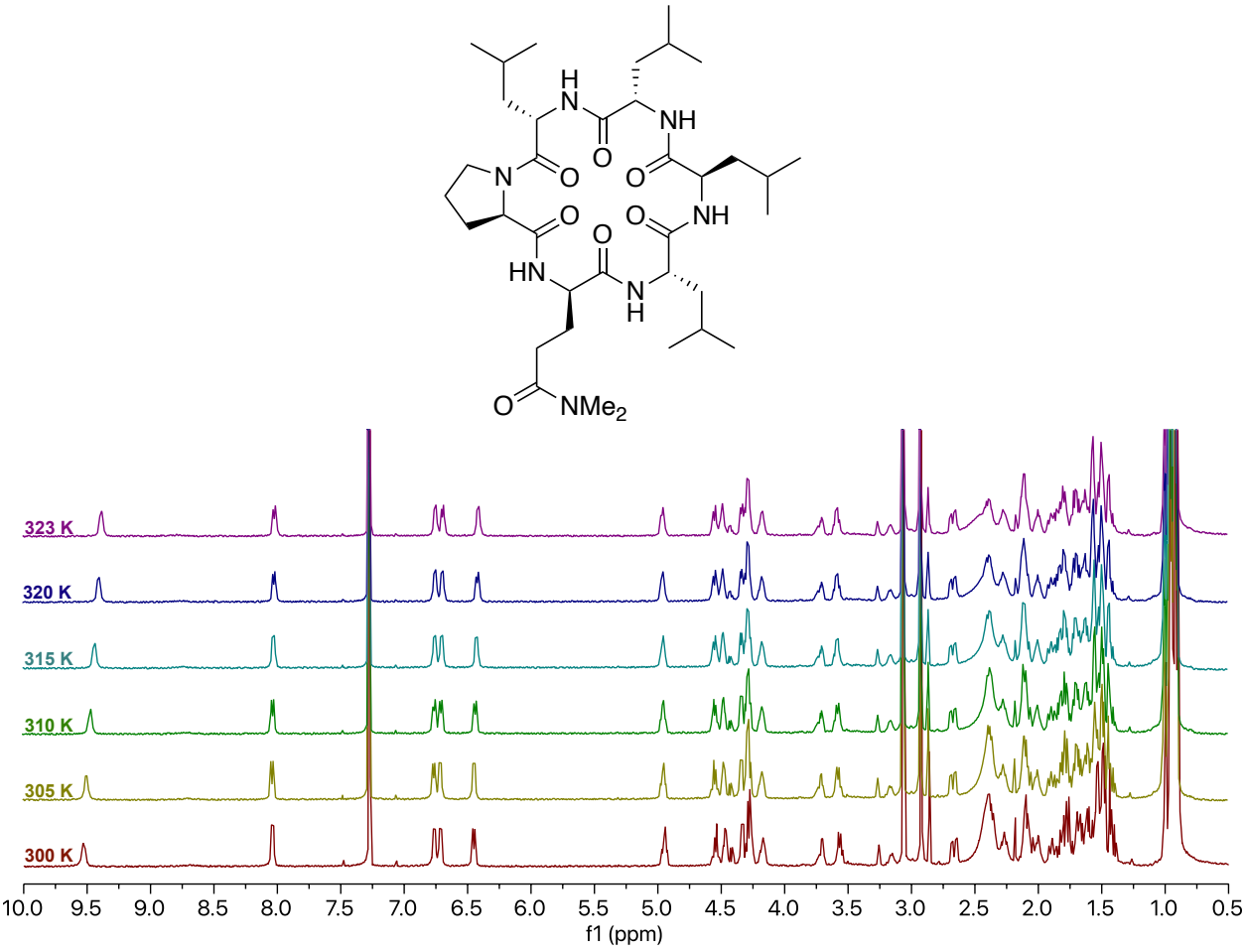


Figure S7. ^{1}H NMR of **3-Glu(OMe)^{2}**.



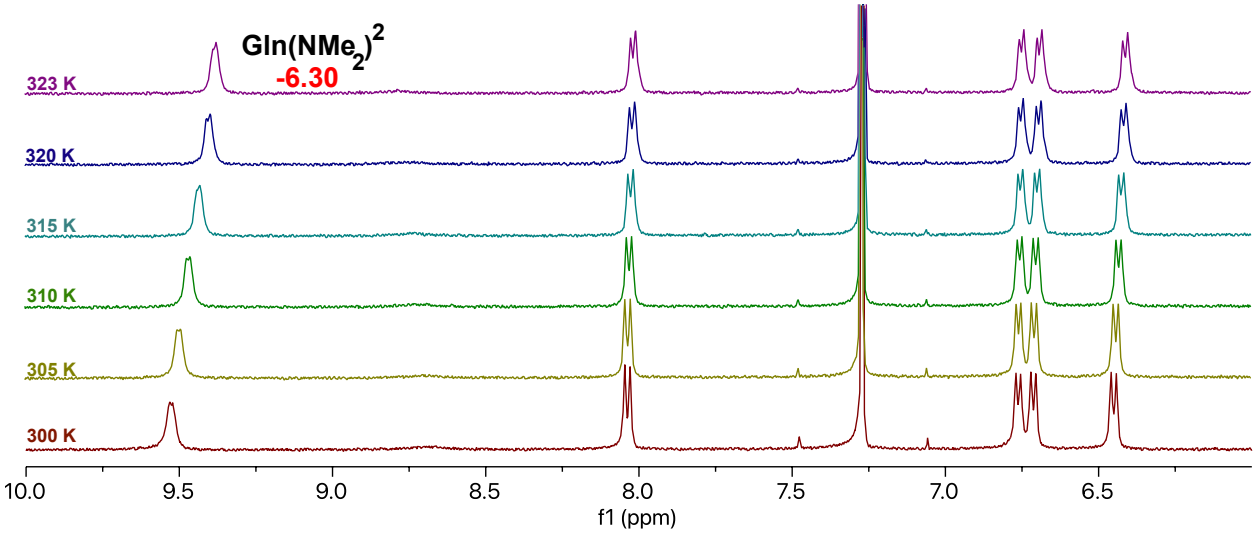
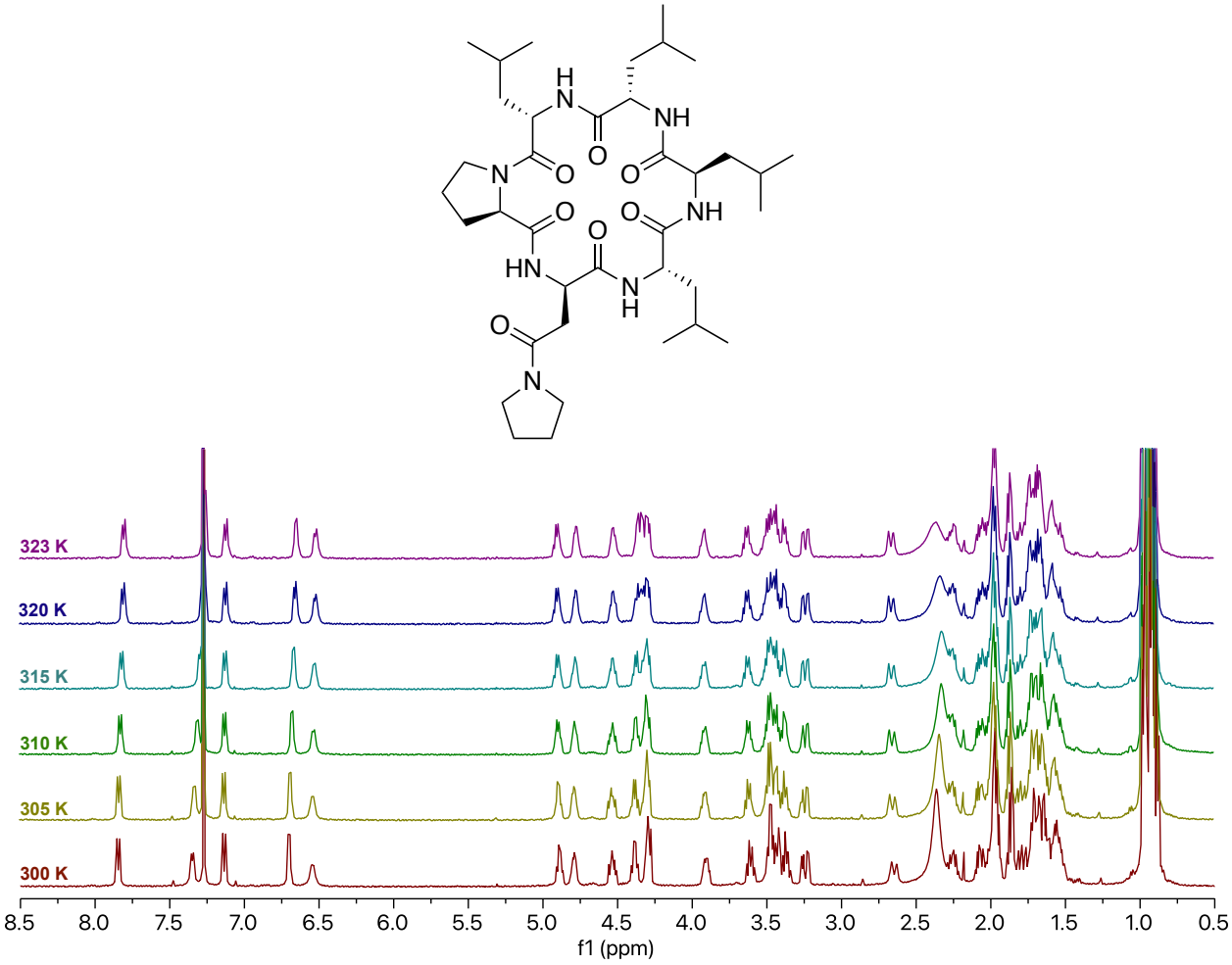


Figure S8. ^{1}H NMR of **3-Glu(NMe$_2$)^2**.



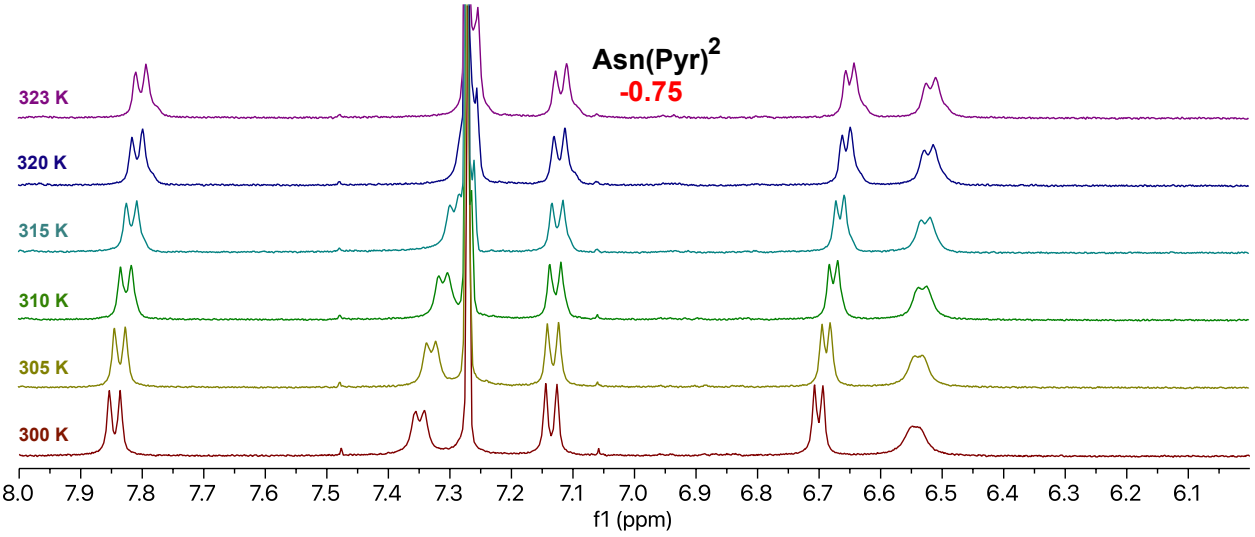
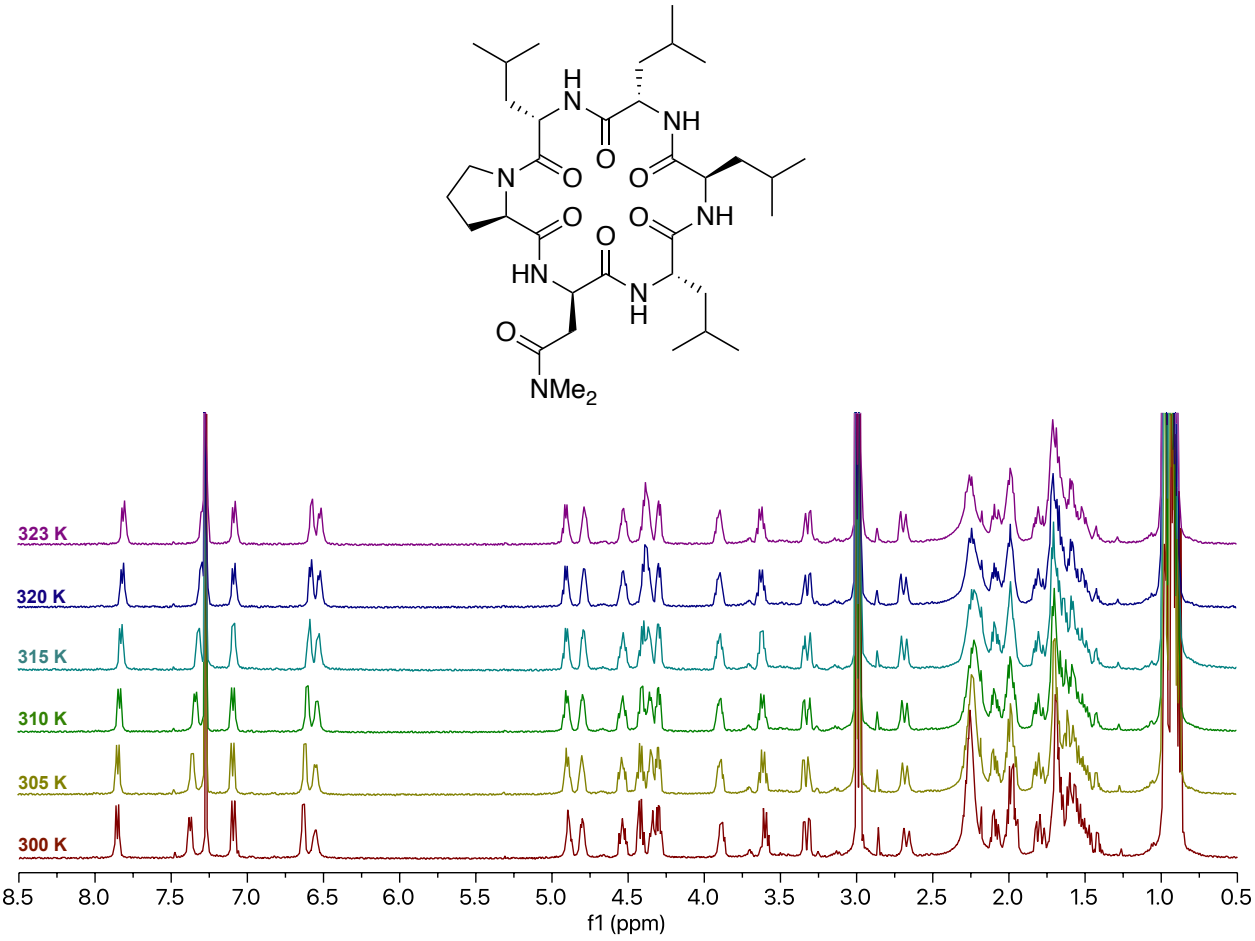


Figure S9. ^{1}H NMR of **3-Asn(Pyr)^{2}**.



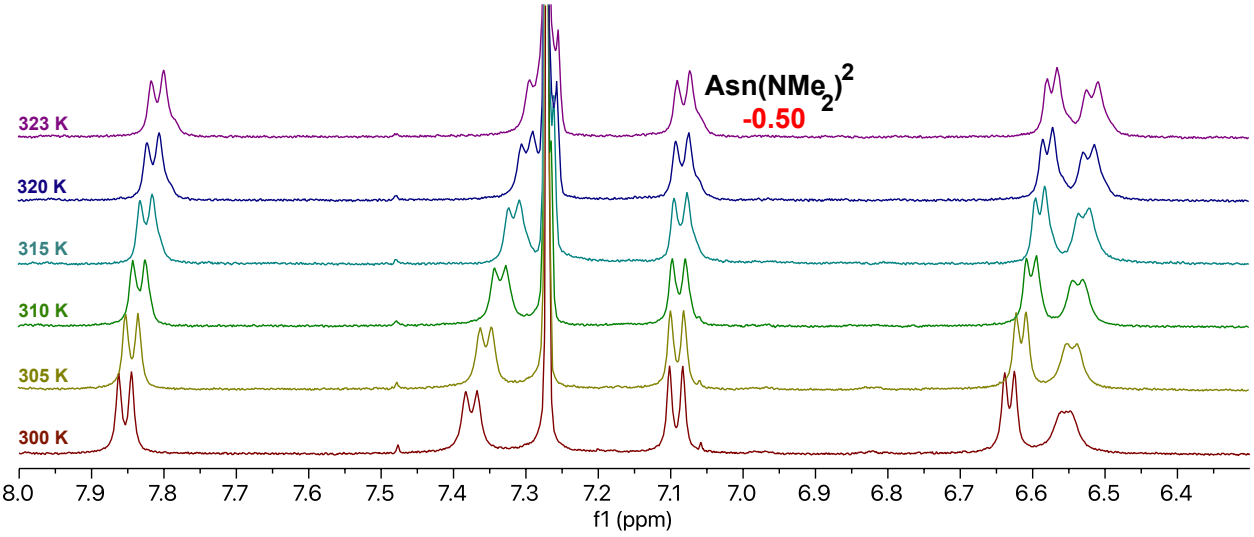


Figure S10. ^{1}H NMR of **3-Asn(NMe$_2$)^2**.

NOE Buildups and Interproton Distances

Interproton distance measurement was done by initial rate approximation.[2-5] ^{1}H-^{1}H NOESY spectra were recorded without solvent suppression with alternated mixing times of 200, 300, 400, 500, 600, 700 and 800 ms. The relaxation delay was set to 2 s. 16 scans were recorded with 2048 points in the direct dimension (F2) and 512 points in the indirect dimension (F1). NOE intensities were calculated by normalization of the integrals of both cross peaks and diagonal peaks of protons *a* and *b*.

$$\mathrm{NOE} = \left[\frac{(\text{cross peak } a * \text{cross peak } b)}{(\text{diagonal peak } a * \text{diagonal peak } b)}\right]^{1/2}$$

Seven normalized NOE intensities were obtained from the different mixing times. Slope from four consecutive mixing time of normalized NOE intensities giving a linear regression ($R^2 \geq 0.95$) was designated as the buildup rates (σ_{ab}).

$$r_{ab} = r_{ref}\left(\frac{\sigma_{ref}}{\sigma_{ab}}\right)^{1/6}$$

The distance between protons *a* and *b* (in Ångström) is indicated by r_{ab} while r_{ref} is the reference distance derived from δ-proline geminal protons (1.78 Å). The buildup rate of the reference protons is indicated by σ_{ref} and σ_{ab} is the buildup rate of protons *a* and *b*. The 10% deviation of calculated distances was applied to account for any errors from NOESY records and calculations. The summary of distance ranges comparing to the distance from crystal structure was summarized in Table S4.

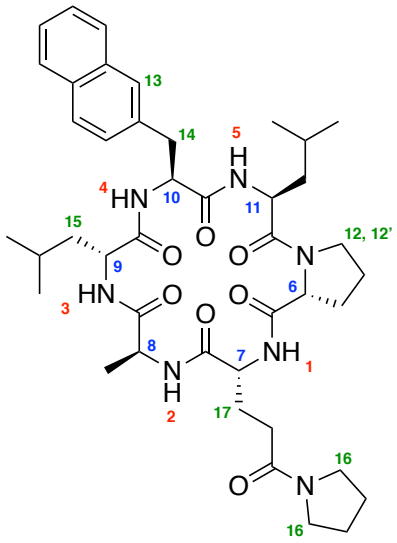
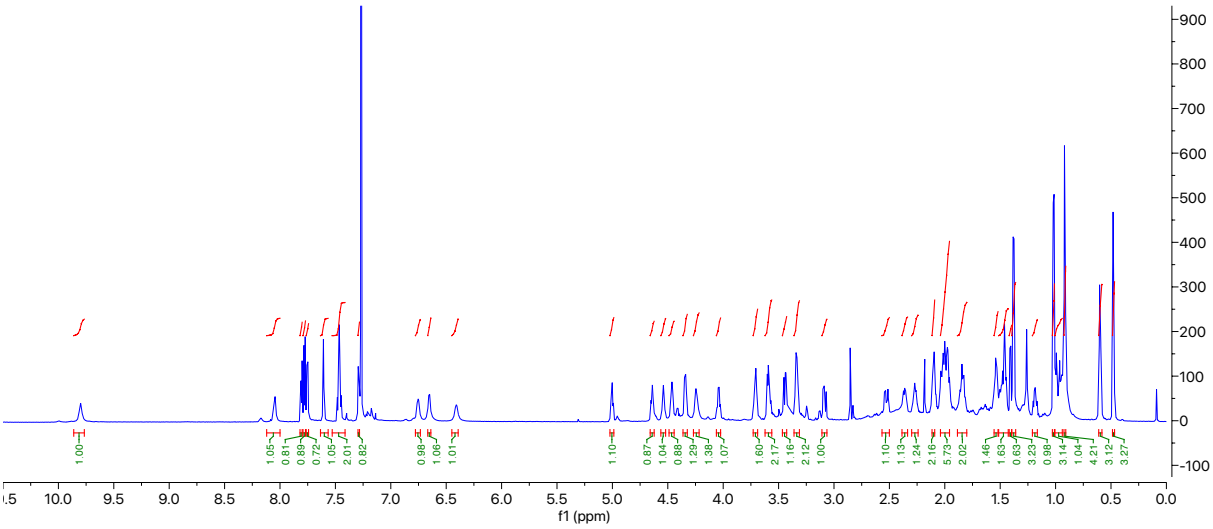


Table S5. Interproton Distances of 3-Pye[2](Ala[3]Nal[5]) Derived from NOE Buildup Rate in $CDCl_3$ and from Crystal Structure

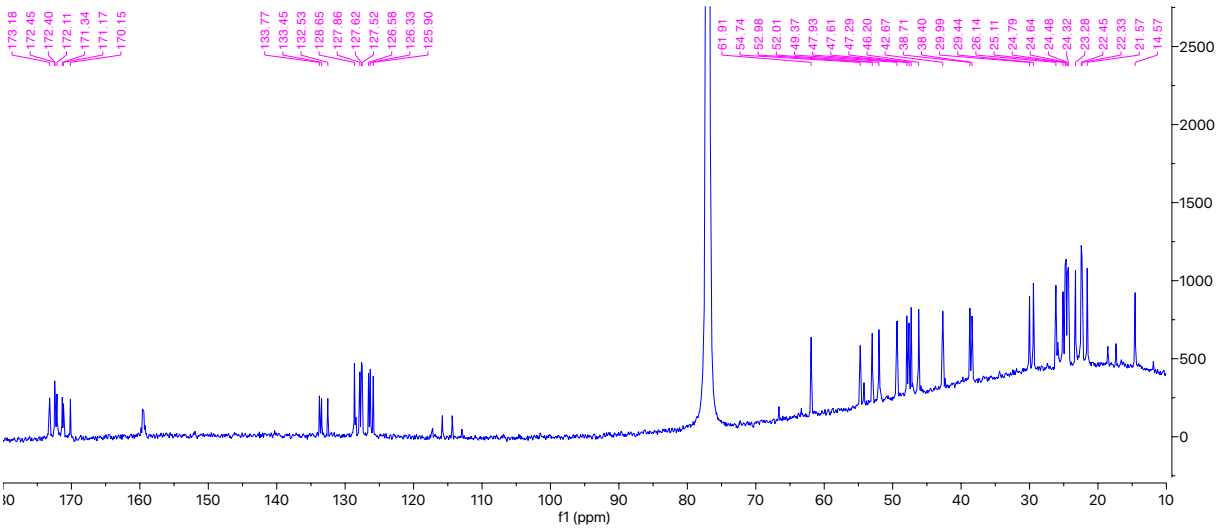
No.	Proton *a*	Proton *b*	$\delta^{1}H_a$	$\delta^{1}H_b$	σ (x 10^{-5})	R^2	r_{ab} [Å]	r_{ab} range (± 10%)	Crys. r_{ab}	Cluster 4 466 conformers		T = 300K ensemble 5000 conformers	
										Avg. r_{ab}	Avg. violation	Avg. r_{ab}	Avg. violation
1	1	2	9.81	8.06	1.47	0.98	2.39	2.15 – 2.63	2.66	2.48	-0.15	2.78	0.15
2	1	12	9.81	4.24	0.65	1.00	2.74	2.46 – 3.01	2.95	3.52	0.51	3.47	0.46
3	1	7	9.81	4.48	0.48	0.99	2.88	2.60 – 3.17	2.65	2.91	-0.26	2.82	-0.35
4	2	5	8.06	6.77	0.21	0.97	3.30	2.97 – 3.63	3.23	3.42	-0.21	3.83	0.20
5	2	8	8.06	4.56	0.50	0.97	2.86	2.58 – 3.15	2.78	2.92	-0.23	2.92	-0.23
6	13	10	7.62	4.66	0.62	0.99	2.76	2.48 – 3.03	2.88	3.44	0.41	3.09	0.06
7	13	14	7.62	3.46	0.98	1.00	2.56	2.30 – 2.81	2.62	2.98	0.17	2.96	0.15
8	5	4	6.76	6.42	1.72	1.00	2.33	2.09 – 2.56	2.54	2.07	-0.49	2.60	0.04
9	5	11	6.76	5.02	0.68	0.98	2.72	2.45 – 2.99	2.79	2.94	-0.05	2.87	-0.12
10	3	15	6.66	1.19	1.48	0.99	2.38	2.15 – 2.62	2.63	2.54	-0.08	2.76	0.14
11	3	8	6.66	4.56	2.00	1.00	2.27	2.04 – 2.50	2.14	2.50	0.00	2.86	0.36
12	4	14	6.42	3.10	1.23	0.98	2.46	2.21 – 2.71	2.42	3.27	0.56	2.94	0.23
13	4	9	6.42	4.06	4.01	0.98	2.02	1.82 – 2.22	2.10	2.45	0.23	2.24	0.02
14	4	10	6.42	4.66	1.00	0.99	2.55	2.29 – 2.80	2.73	2.85	0.05	2.84	0.04
15	11	12	5.02	4.26	2.88	0.97	2.13	1.92 – 2.35	2.14	2.33	-0.02	2.28	-0.07
16	11	12'	5.02	3.61	1.65	1.00	2.34	2.11 – 2.58	2.95	2.35	-0.23	2.55	-0.03
17	10	16	4.66	3.60	0.47	0.95	2.89	2.60 – 3.18	2.95	4.03	0.85	4.21	1.03
18	7	17	4.48	2.38	2.15	1.00	2.24	2.02 – 2.47	2.30	2.27	-0.20	2.27	-0.20
19	9	15	4.05	1.02	1.83	0.97	2.30	2.07 – 2.53	2.33	2.63	0.10	2.62	0.09
Ref.	12	12'	4.26	3.61	8.58	1.00	1.78						
										stdev	0.33		0.31

^{1}H NMR



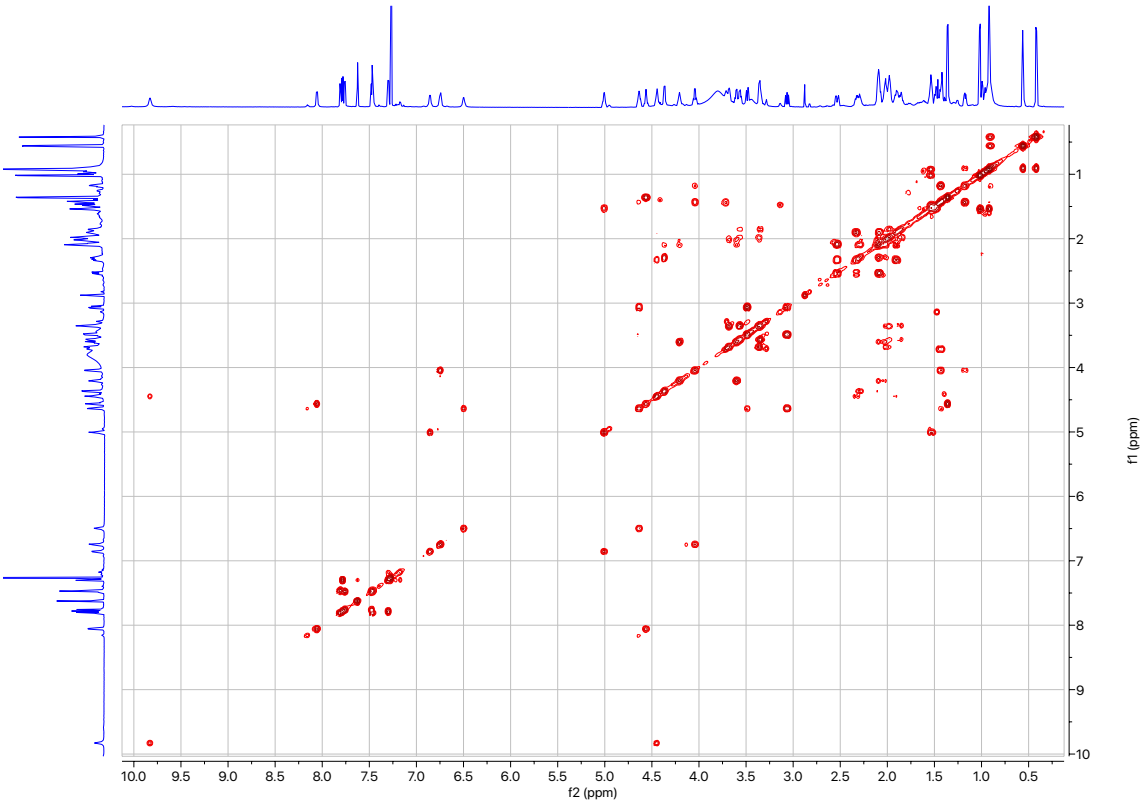
^{1}H NMR (800 MHz, $CDCl_3$) δ 9.80 (br, 1H), 8.05 (d, *J* = 8.7 Hz, 1H), 7.80 (d, *J* = 7.6 Hz, 1H), 7.78 (d, *J* = 8.4 Hz, 1H), 7.75 (d, *J* = 7.0 Hz, 1H), 7.61 (s, 1H), 7.49 – 7.44 (m, 2H), 7.29 (d, *J* = 9.4 Hz, 1H), 6.75 (br, 1H), 6.65 (d, *J* = 7.9 Hz, 1H), 6.41 (br, 1H), 5.00 (td, *J* = 8.1, 4.8 Hz, 1H), 4.64 (td, *J* = 9.2, 4.5 Hz, 1H), 4.54 (p, *J* = 7.0 Hz, 1H), 4.49 – 4.44 (m, 1H), 4.34 (dd, *J* = 9.0, 3.8 Hz, 1H), 4.27 – 4.22 (m, 1H), 4.04 (q, *J* = 8.0 Hz, 1H), 3.74 – 3.68 (m, 1H), 3.62 – 3.56 (m, 2H), 3.45 (dd, *J* = 14.2, 4.5 Hz, 1H), 3.36 – 3.31 (m, 2H), 3.09 (dd, *J* = 14.2, 9.6 Hz, 1H), 2.53 (dd, *J* = 17.1, 8.2 Hz, 1H), 2.39 – 2.34 (m, 1H), 2.30 – 2.24 (m, 1H), 2.13 – 2.07 (m, 2H), 2.06 – 1.94 (m, 5H), 1.85 (m, 2H), 1.54 – 1.51 (m, 1H), 1.51 – 1.43 (m, 2H), 1.41 (d, *J* = 6.7 Hz, 1H), 1.38 (d, *J* = 6.9 Hz, 3H), 1.21 – 1.16 (m, 1H), 1.02 (d, *J* = 6.0 Hz, 3H), 1.01 – 0.94 (m, 1H), 0.92 (d, *J* = 6.3 Hz, 3H), 0.60 (d, *J* = 6.6 Hz, 3H), 0.48 (d, *J* = 6.6 Hz, 3H).

^{13}C NMR

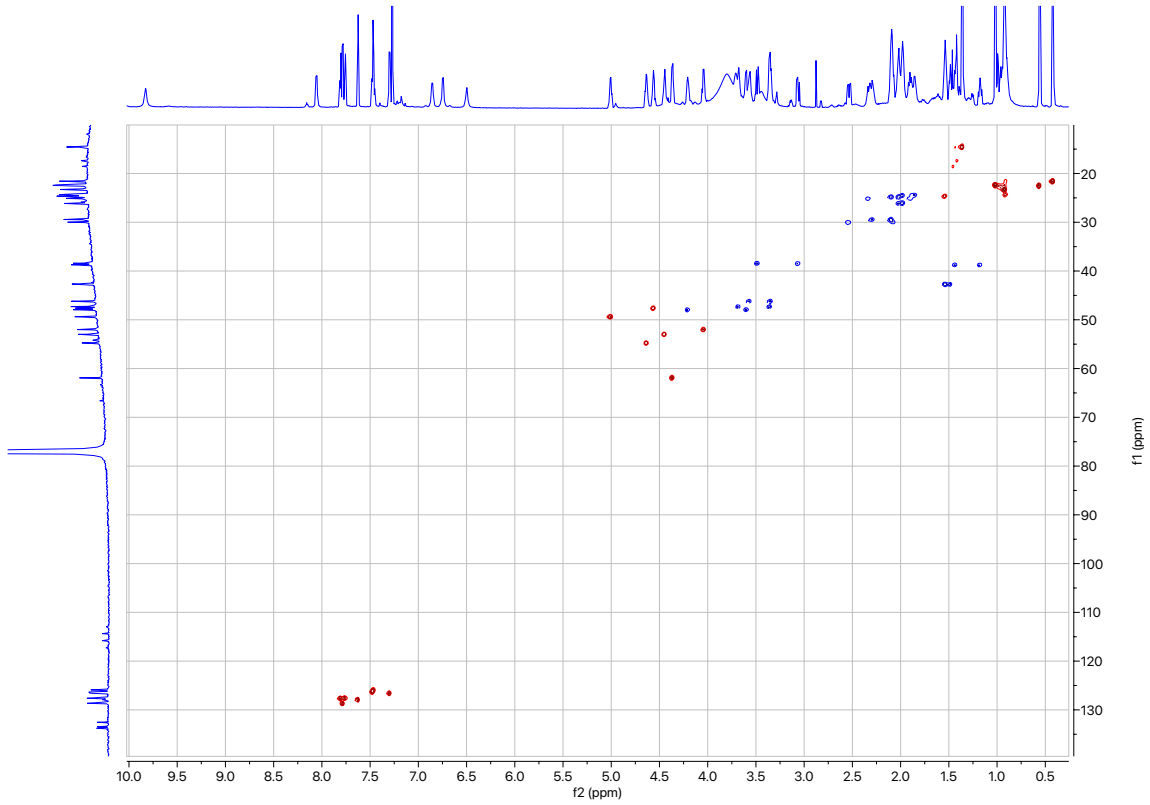


^{13}C NMR (201 MHz, $CDCl_3$) δ 173.18, 172.45, 172.40, 172.11, 171.34, 171.17, 170.15, 133.77, 133.45, 132.53, 128.65, 127.86, 127.62, 127.52, 126.58, 126.33, 125.90, 61.91, 54.74, 52.98, 52.01, 49.37, 47.93, 47.61, 47.29, 46.20, 42.67, 38.71, 38.40, 29.99, 29.44, 26.14, 25.11, 24.79, 24.64, 24.48, 24.32, 23.28, 22.45, 22.33, 21.57, 14.57.

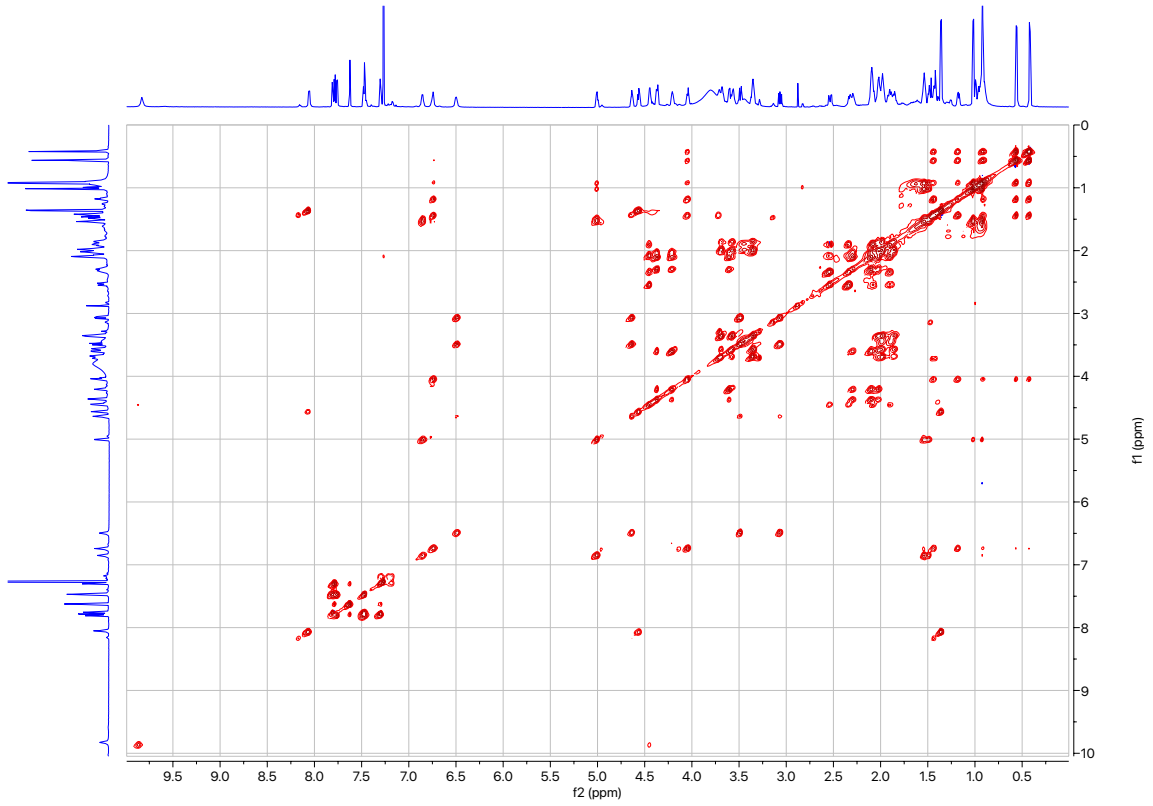
^{1}H-^{1}H COSY



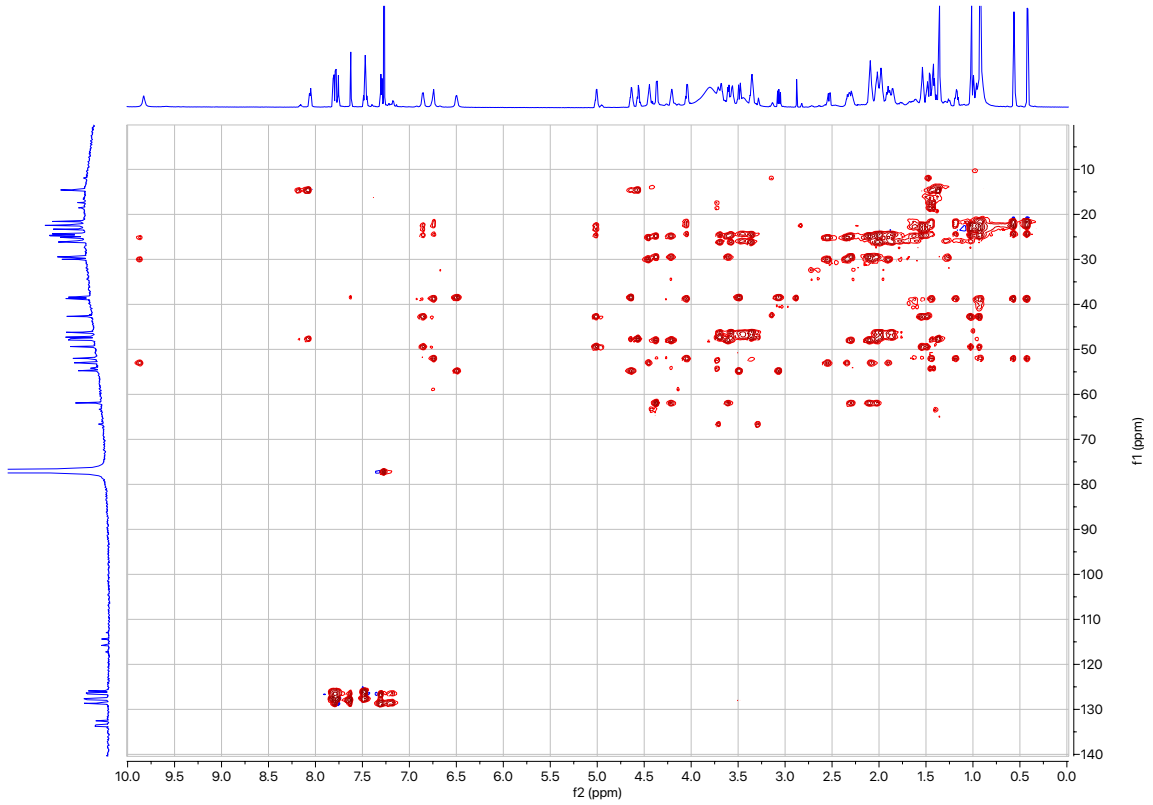
^{1}H-^{13}C HSQC



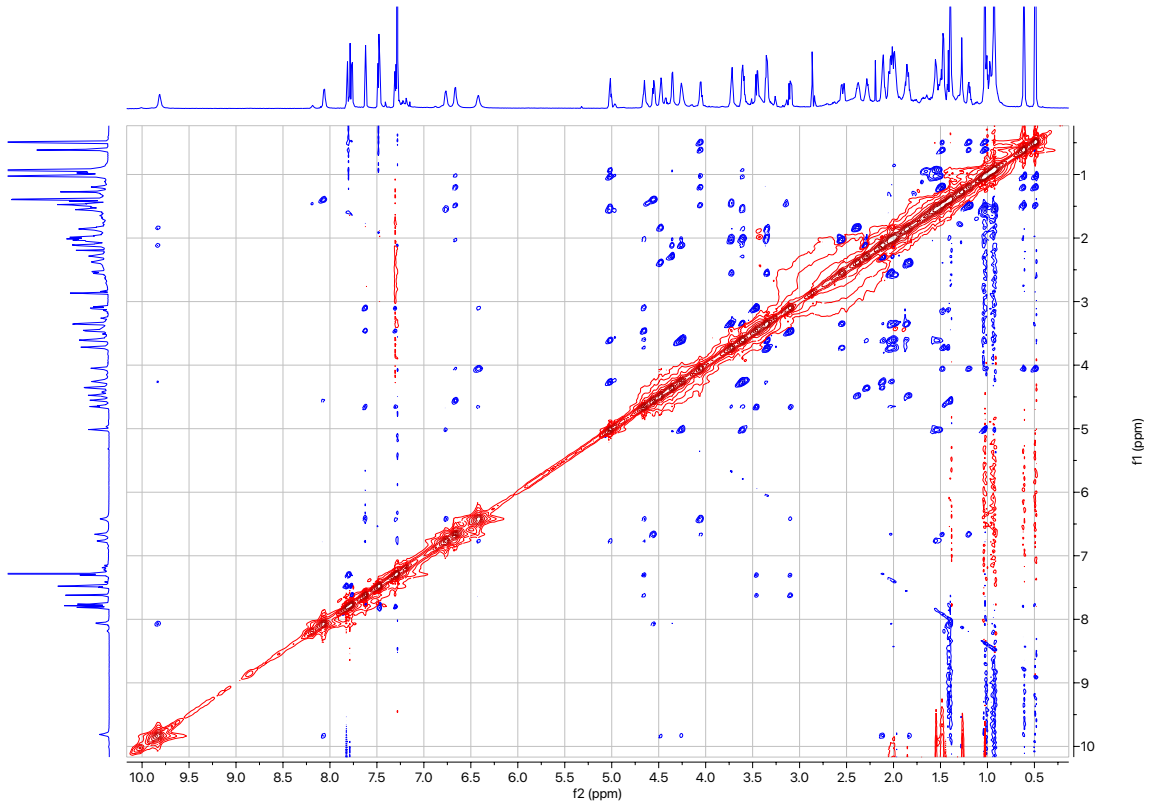
^{1}H-^{1}H TOCSY



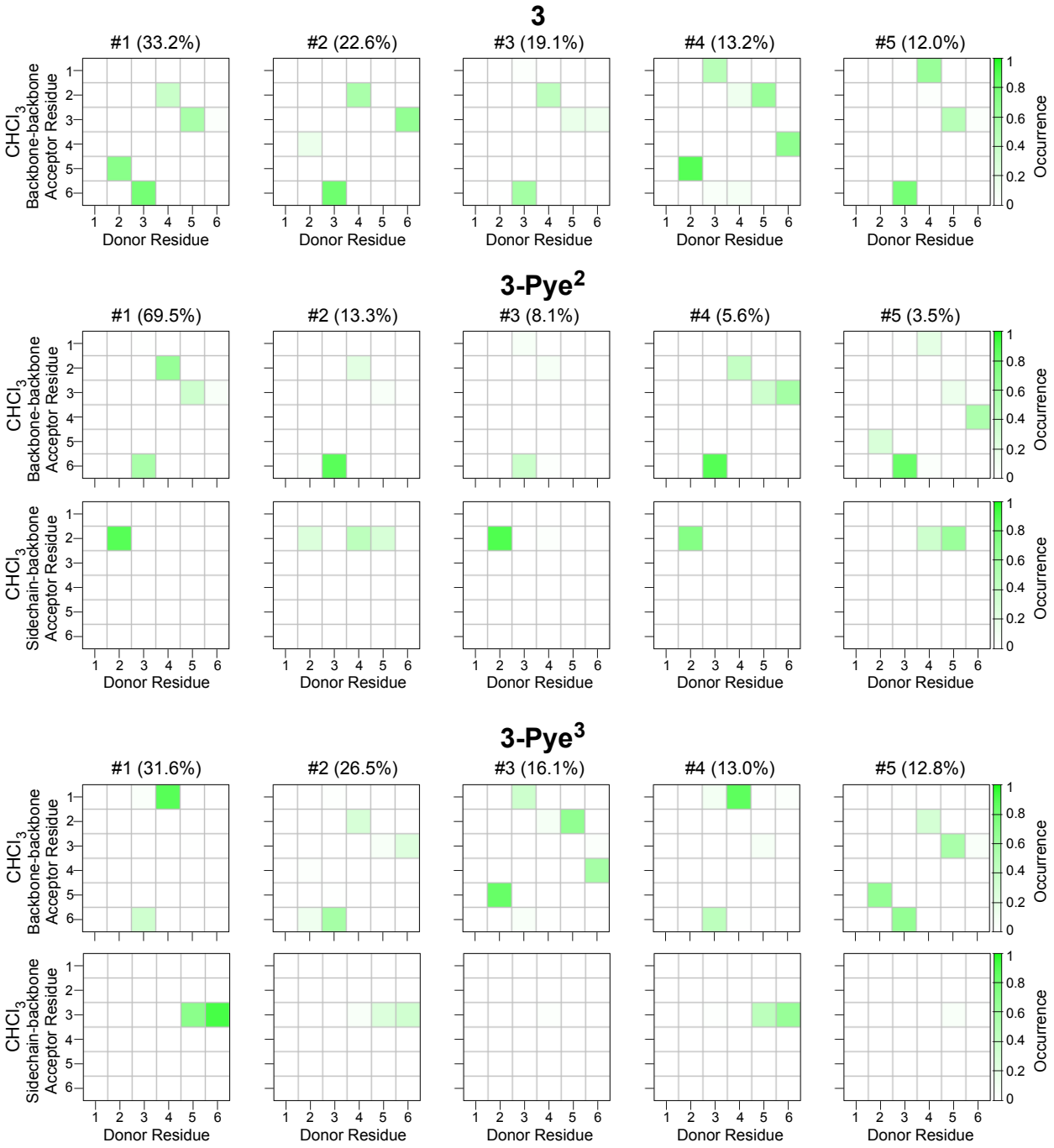
^{1}H-^{13}C HSQC-TOCSY

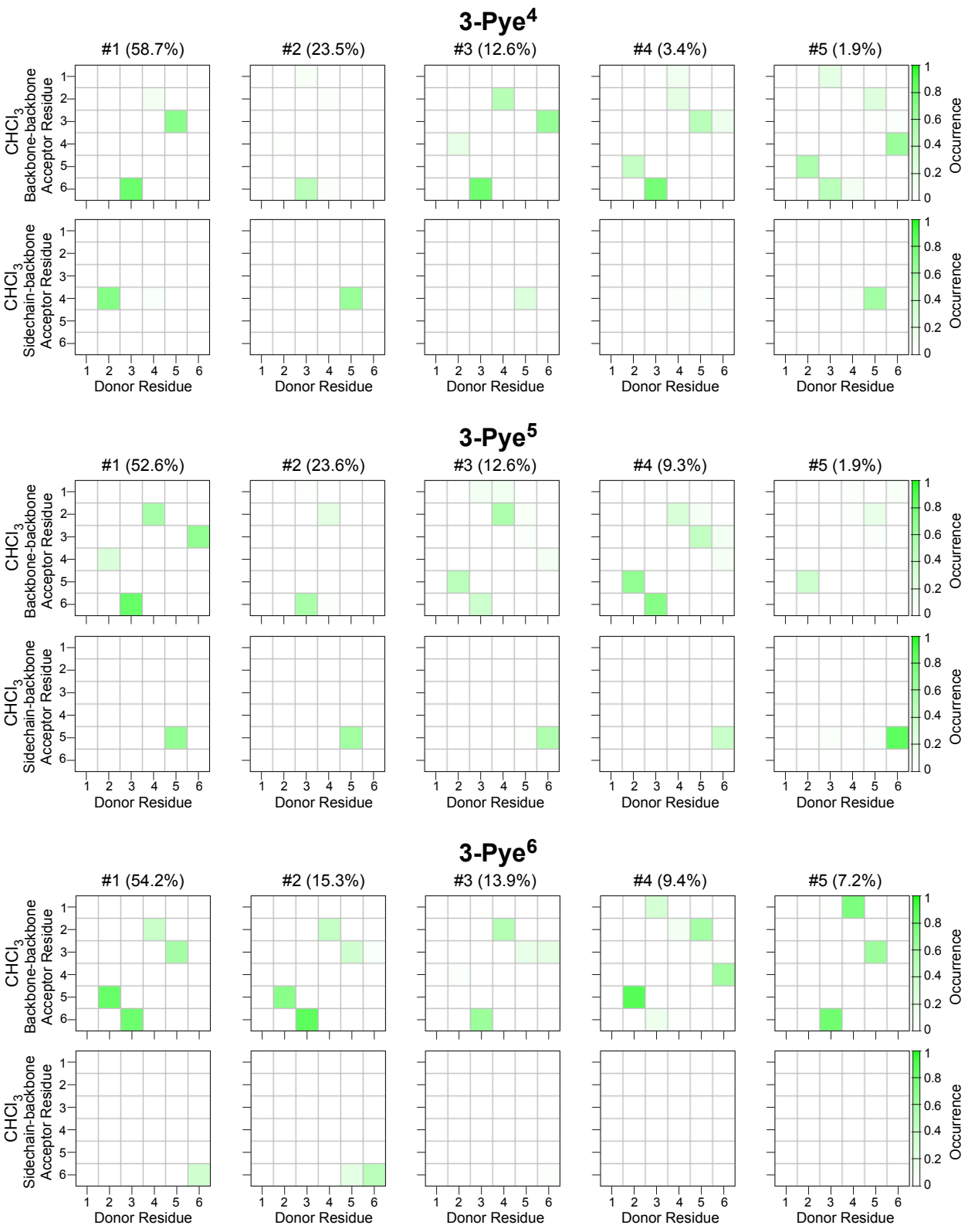


^{1}H-^{1}H NOESY


10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f2 (ppm)
1
2
3
4
5
6
7
8
9
10
f1 (ppm)

Intramolecular Hydrogen Bonding (IMHB) Patterns




3-Pye4
#1 (58.7%)
#2 (23.5%)
#3 (12.6%)
#4 (3.4%)
#5 (1.9%)
3-Pye5
#1 (52.6%)
#2 (23.6%)
#3 (12.6%)
#4 (9.3%)
#5 (1.9%)
3-Pye6
#1 (54.2%)
#2 (15.3%)
#3 (13.9%)
#4 (9.4%)
#5 (7.2%)
CHCl3 Backbone-backbone Acceptor Residue
CHCl3 Sidechain-backbone Acceptor Residue
Donor Residue
Occurrence

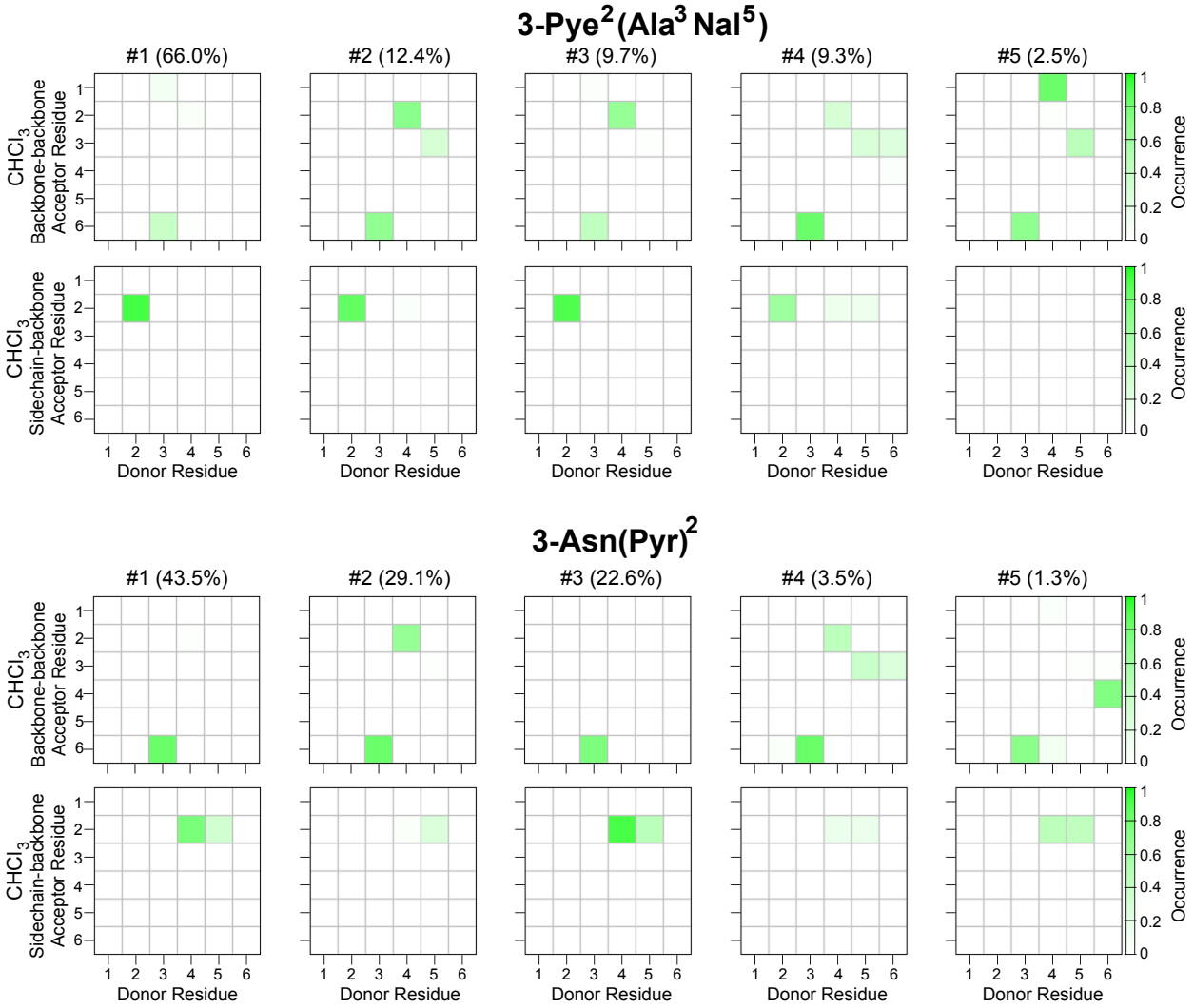


Figure S11 IMHB patterns of compound **3** series.

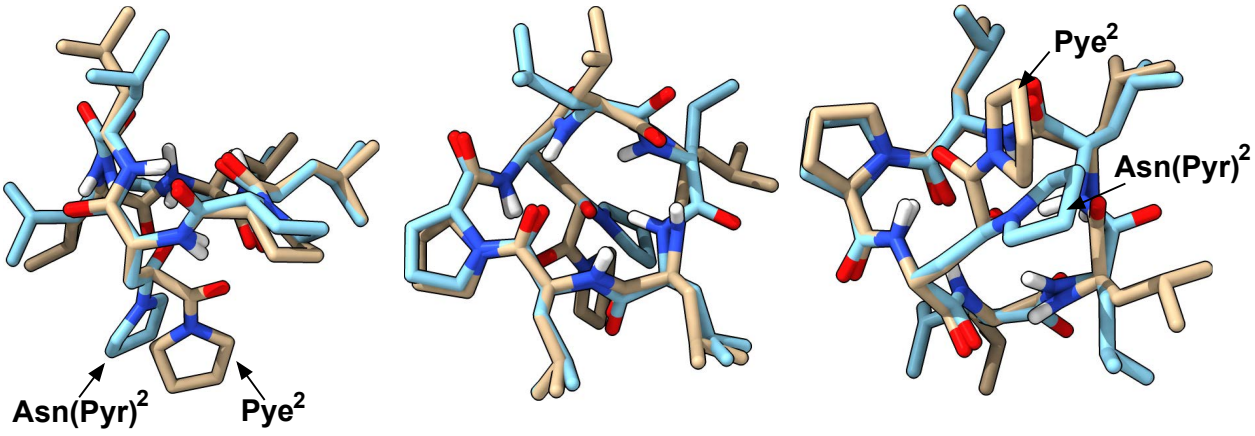


Figure S12 Overlay between **3-Pye²** (gold) and **3-Asn(Pyr)²** (blue). Backbone rmsd = 0.72 Å

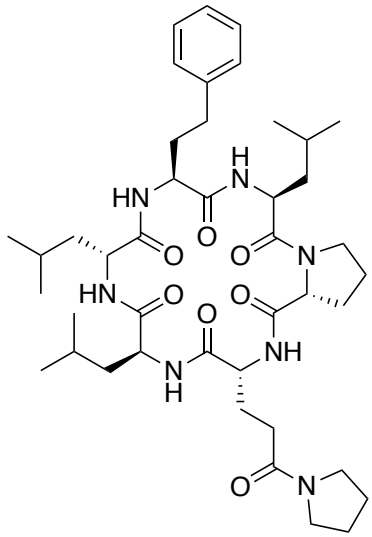
General Procedure of X-ray Structure Determination

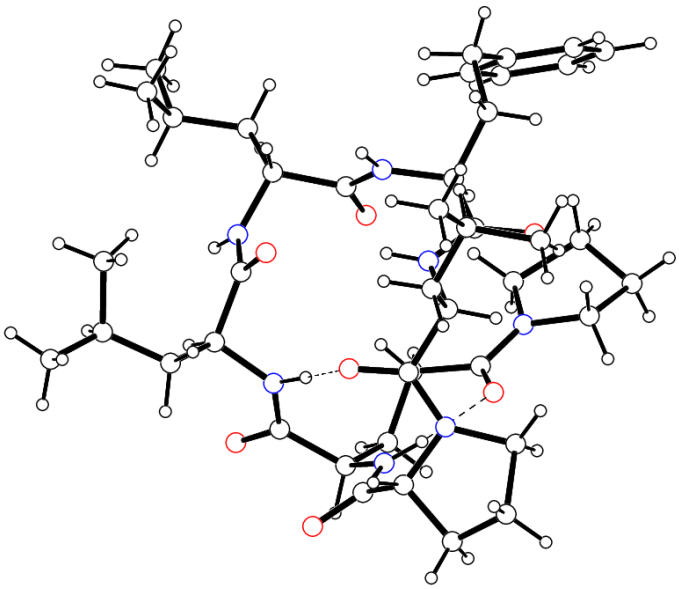
Crystals of **3-Pye^2(Ala^{3}Nal^5)** and **3-Pye^2(HPhe^5)** were obtained by pentane-THF vapor diffusion. In each case, a prism was microscopically selected under crossed polarizers, mounted on a MiTeGen polyimide loop, and cooled to 100 K on a Rigaku Synergy-S X-ray diffractometer. Diffraction of Cu Kα radiation from a PhotonJet-S microfocus source was detected using a HyPix-6000HE hybrid photon counting detector. Bijvoet pairs were collected for absolute structure determination. Screening, indexing, data collection, and data processing were performed with CrysAlis^{Pro}.[6] The structure was solved using SHELXT and refined using SHELXL following established strategies.[7-9] All non-H atoms were refined anisotropically. H atoms were placed at calculated positions and refined with a riding model and coupled isotropic displacement parameters (1.5 × U_{eq} for methyl groups and 1.2 × U_{eq} for all others). Anomalous dispersion was used to refine the absolute structure (Flack) parameter.

Crystal Data and Structure Refinement for 3-Pye^2(HPhe^5)

Empirical formula	C42 H65 N7 O7
Formula weight	780.01
Temperature	100.1(1) K
Wavelength	1.54184 Å
Crystal system	Trigonal
Space group	$P3_221$
Unit cell dimensions	a = 18.9888(2) Å
	b = 18.9888(2) Å
	c = 22.4565(2) Å
Volume	7012.41(16) Å^3
Z	6
Density (calculated)	1.108 Mg/m^3
Absorption coefficient	0.614 mm^{-1}
F(000)	2532
Crystal size	0.32 x 0.07 x 0.06 mm^3
Theta range for data collection	2.687 to 70.071°.
Index ranges	-17<=h<=23, -22<=k<=23, -27<=l<=27
Reflections collected	60958
Independent reflections	8894 [R(int) = 0.0306]
Completeness to theta = 67.684°	100.0 %
Absorption correction	Gaussian
Max. and min. transmission	1.000 and 0.763
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	8894 / 1611 / 711
Goodness-of-fit on F^2	1.040

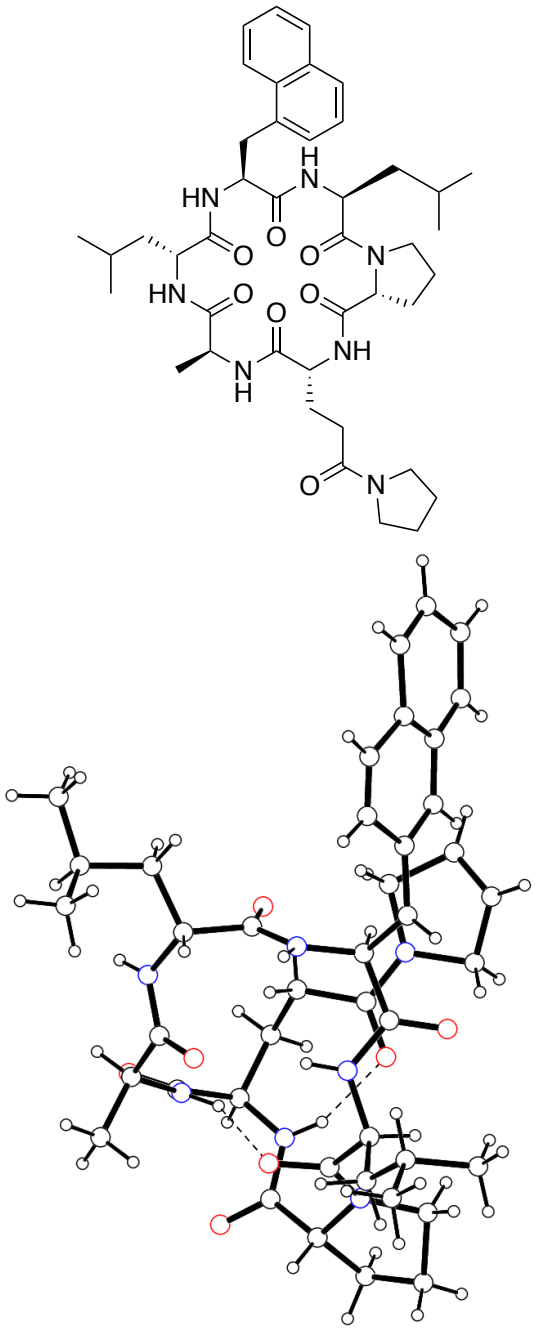
Final R indices [I>2σ(I)]	R1 = 0.0606, wR2 = 0.1734
R indices (all data)	R1 = 0.0637, wR2 = 0.1773
Absolute structure parameter	0.1(3)
Extinction coefficient	n/a
Largest diff. peak and hole	0.400 and -0.252 e.Å^{-3}





Crystal Data and Structure Refinement for 3-Pye^2(Ala^{3}Nal^5)

Empirical formula	C42 H59 N7 O7
Formula weight	773.96
Temperature	100.0(1) K
Wavelength	1.54184 Å
Crystal system	Trigonal
Space group	$P3_221$
Unit cell dimensions	a = 18.8195(2) Å
	b = 18.8195(2) Å
	c = 22.4826(3) Å
Volume	6895.94(17) Å^3
Z	6
Density (calculated)	1.118 Mg/m^3
Absorption coefficient	0.622 mm^{-1}
F(000)	2496
Crystal size	0.13 x 0.07 x 0.05 mm^3
Theta range for data collection	2.711 to 70.067°.
Index ranges	-22<=h<=20, -22<=k<=22, -27<=l<=26
Reflections collected	90045
Independent reflections	8733 [R(int) = 0.0369]
Completeness to theta = 67.684°	100.0 %
Absorption correction	Gaussian
Max. and min. transmission	1.000 and 0.858
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	8733 / 1957 / 806
Goodness-of-fit on F^2	1.044
Final R indices [I>2σ(I)]	R1 = 0.0714, wR2 = 0.2074
R indices (all data)	R1 = 0.0781, wR2 = 0.2172
Absolute structure parameter	0.1(4)
Extinction coefficient	n/a
Largest diff. peak and hole	0.473 and -0.291 e.Å^{-3}



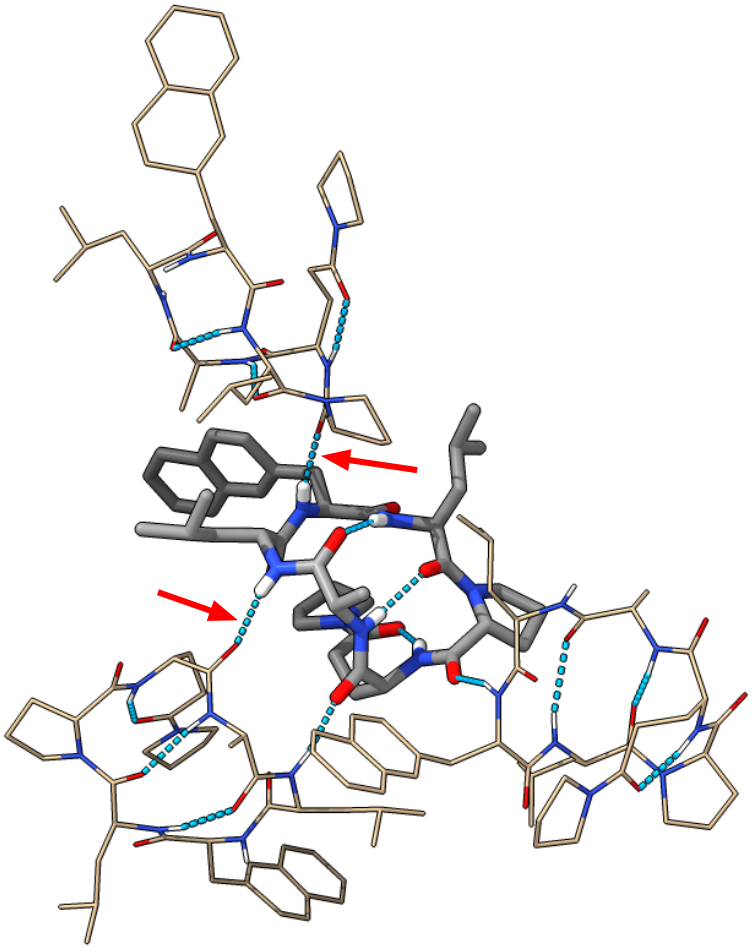


Figure S13. Crystal lattice of **3-Pye^{2}-Ala^{3}-^{2}Nal^{5}** shows intermolecular hydrogen bonds (red arrows).

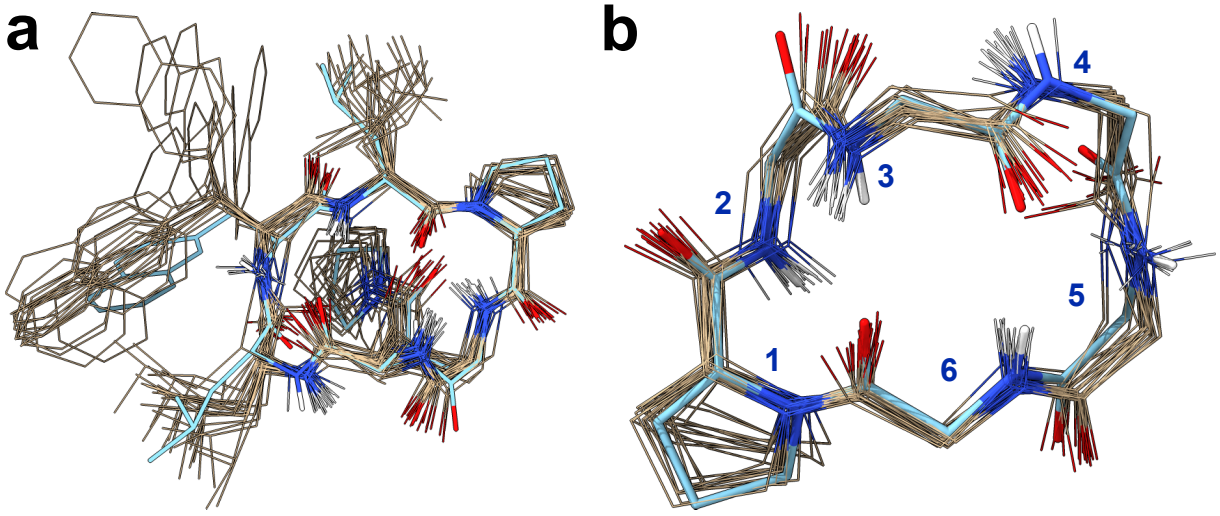


Figure S14. (**a**) Overlays of **3-Pye^{2}(Ala^{3}Nal^{5})** crystal structure (blue) and 20 simulated McMD structures (gold). (**b**) Crystal structure and simulated structures show similar backbone geometry. Amide protons of simulated 3rd and 6th residues are consistent to the crystal structure, while amide 4th and 5th NH distort from crystal structure which allow to form IMHB.

LC-MS Spectra

Column: Kinetex® C18 (30 x 2.1 mm, 2.6 µm 100 Å)

Wavelength: 200 or 254 nm

Temperature: 50 ºC

Flow rate: 0.4 mL/min

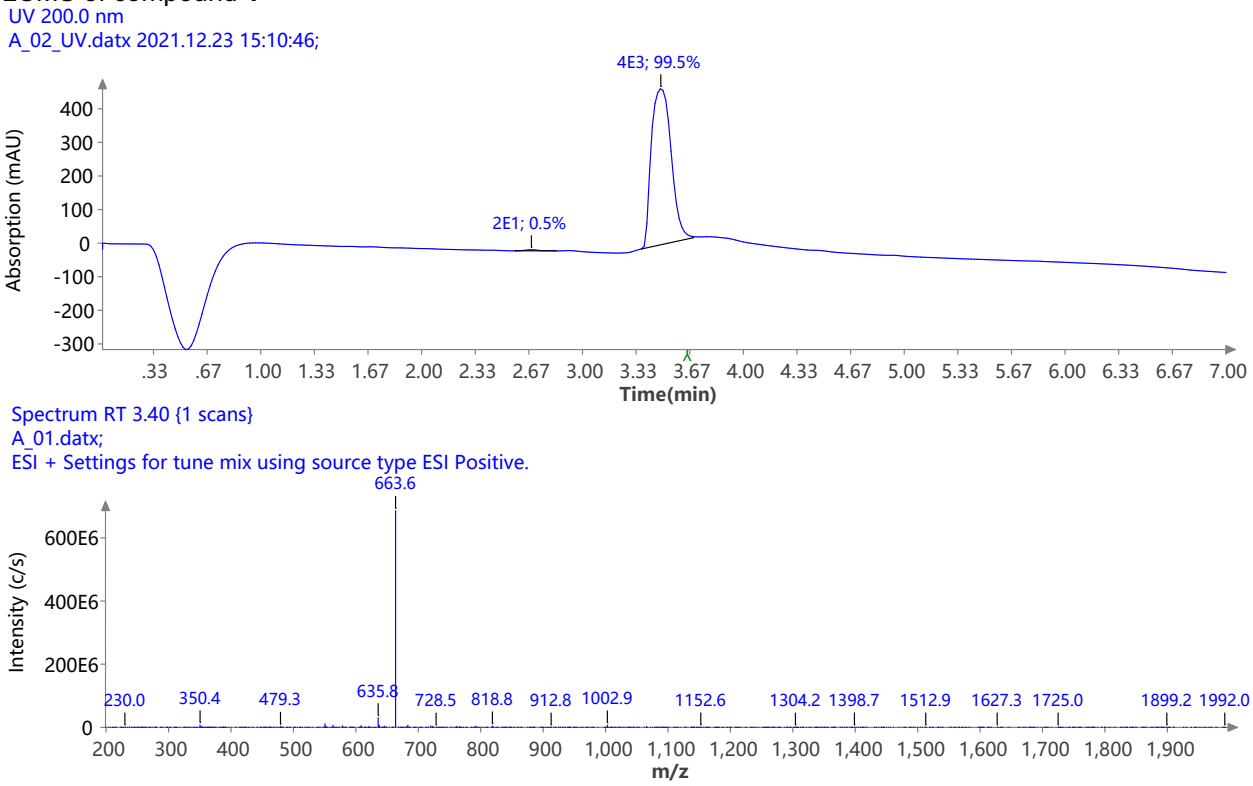
Injection: 10 µL

Mobile phase: A phase: 0.1% formic acid in Milli-Q water

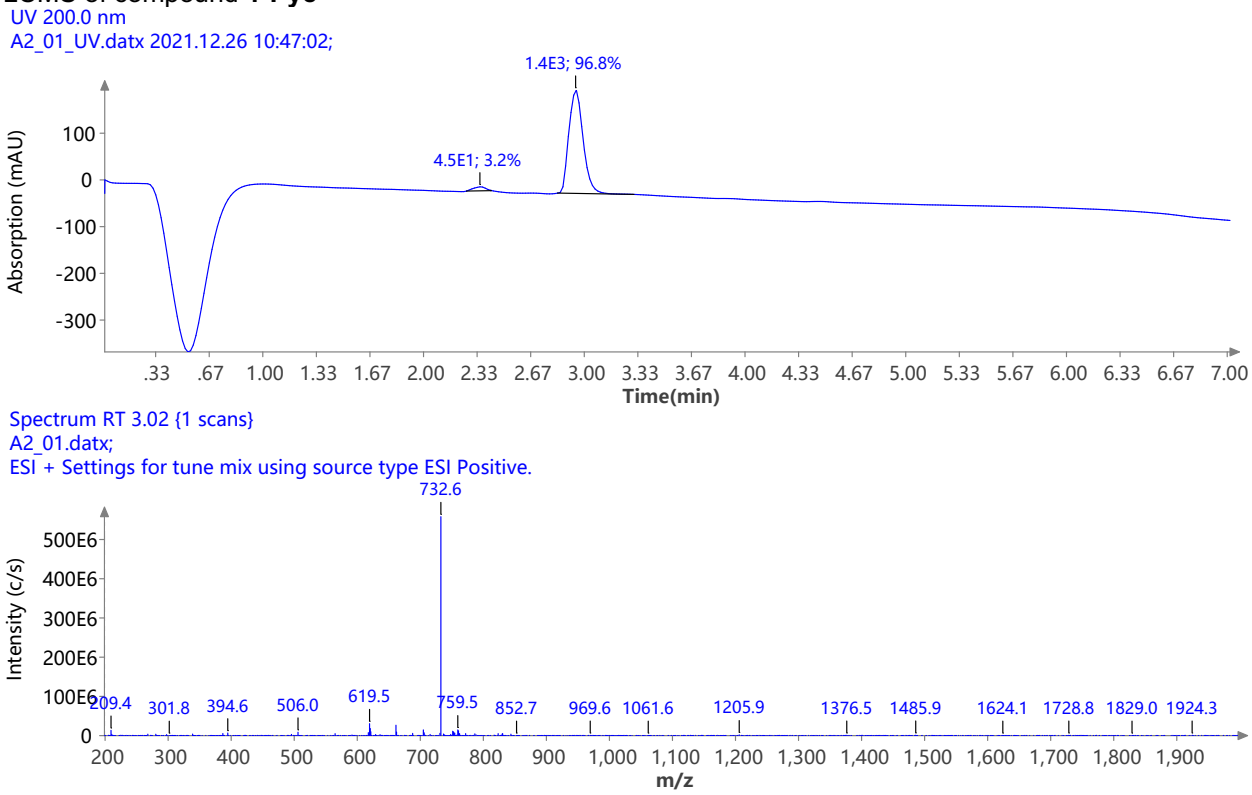
B phase: 0.1% formic acid in CH_3CN

Gradient elution: 0-6 min (20-100% B), 6-7 min (100% B)

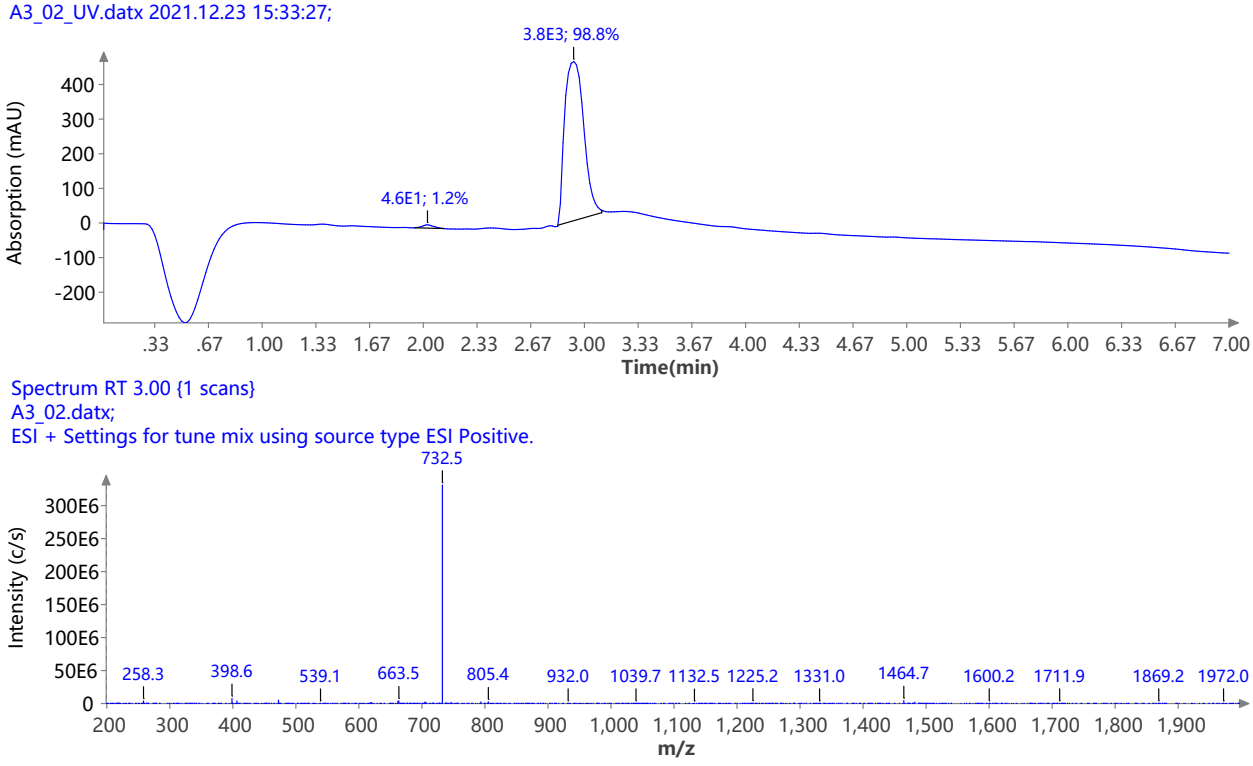
LCMS of compound **1**



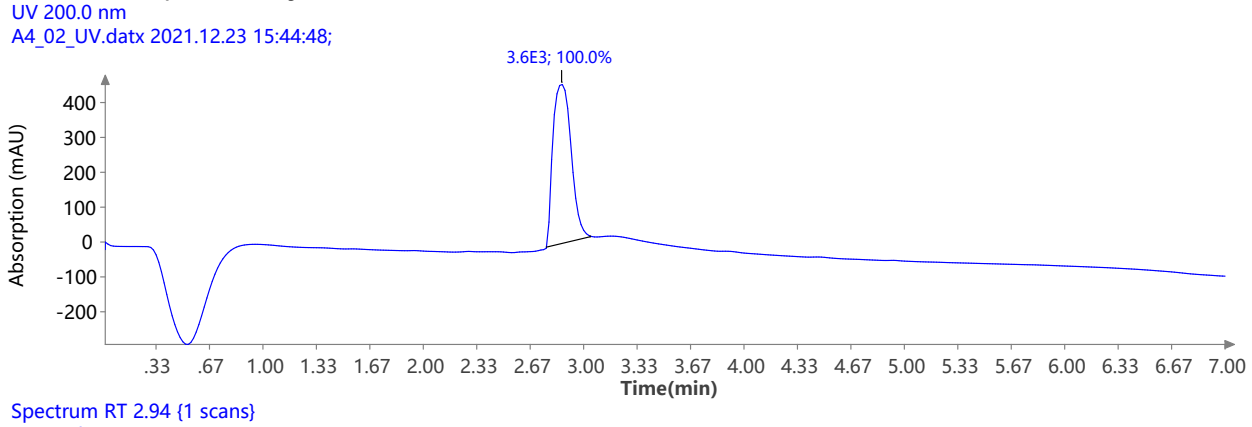
LCMS of compound **1-Pye^2**

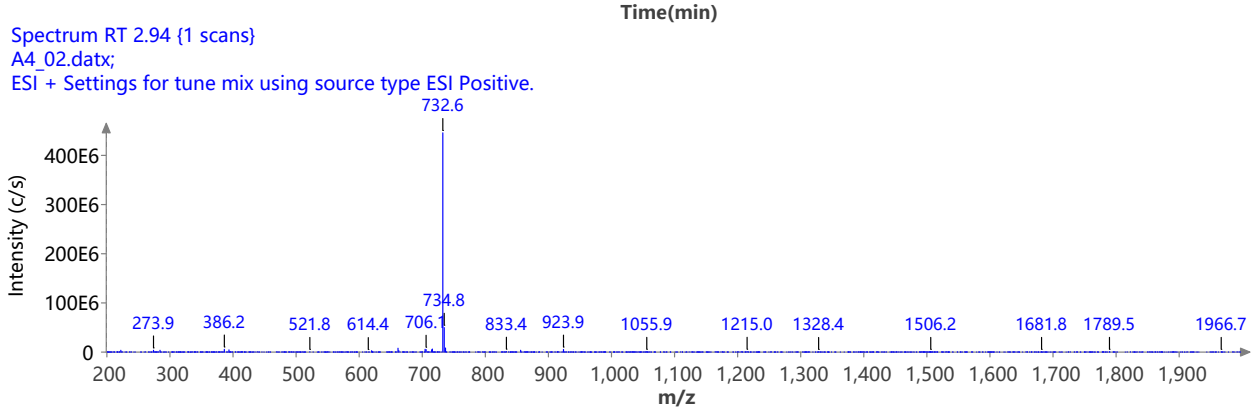


LCMS of compound **1-Pye^3**

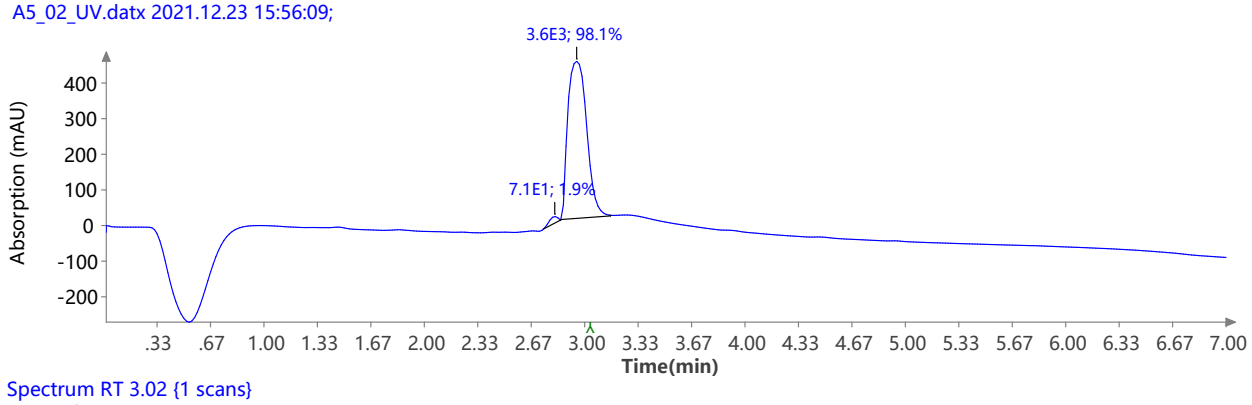


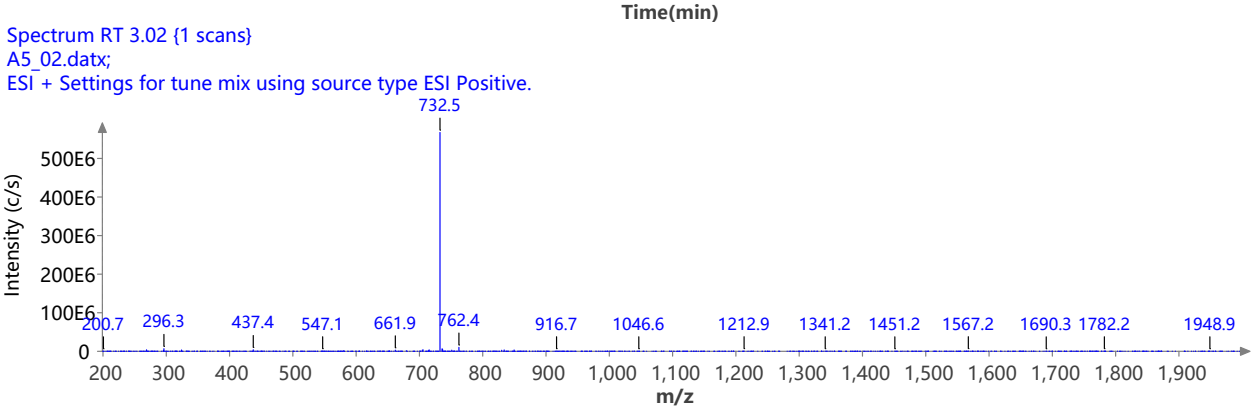
LCMS of compound **1-Pye[4]**



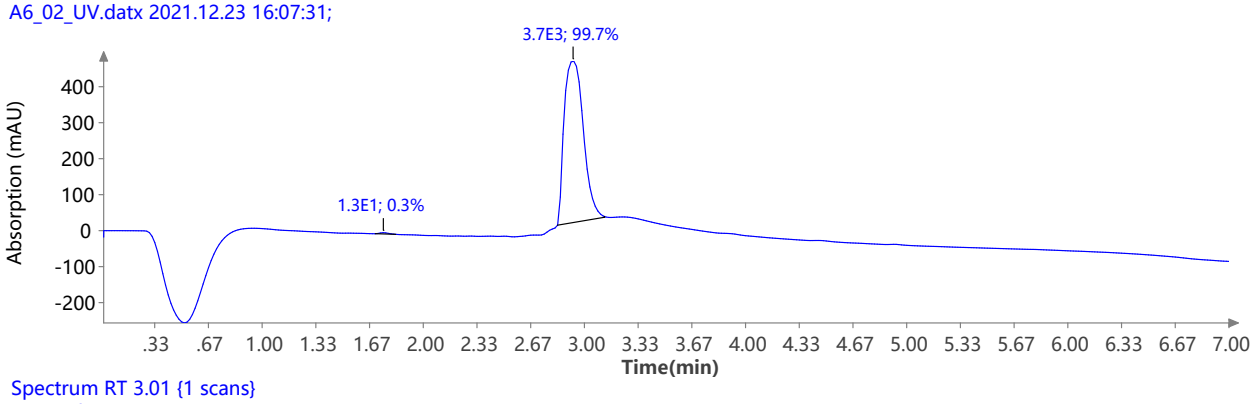


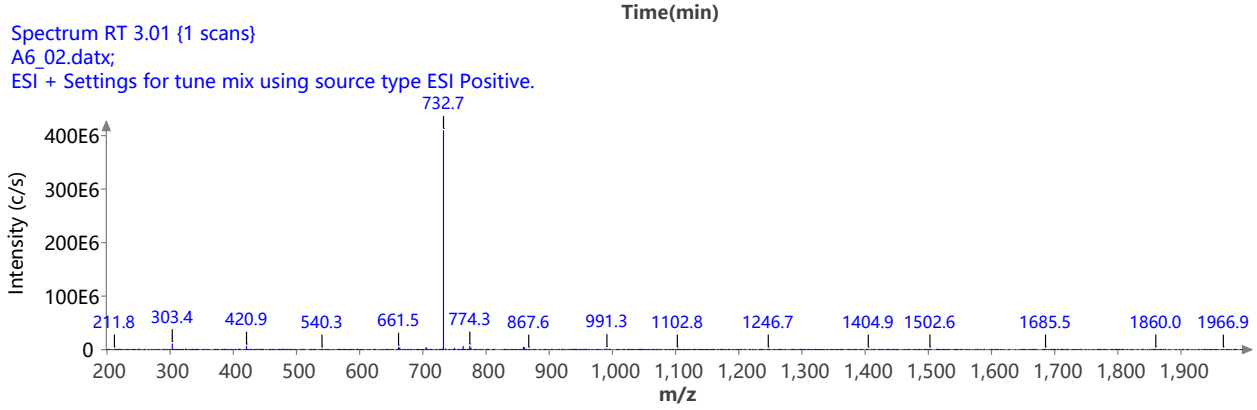
LCMS of compound **1-Pye[5]**



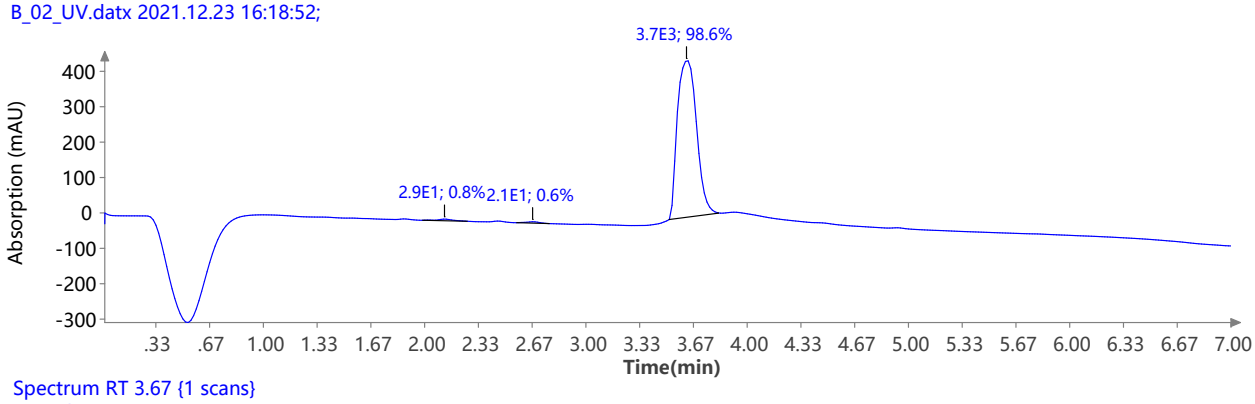


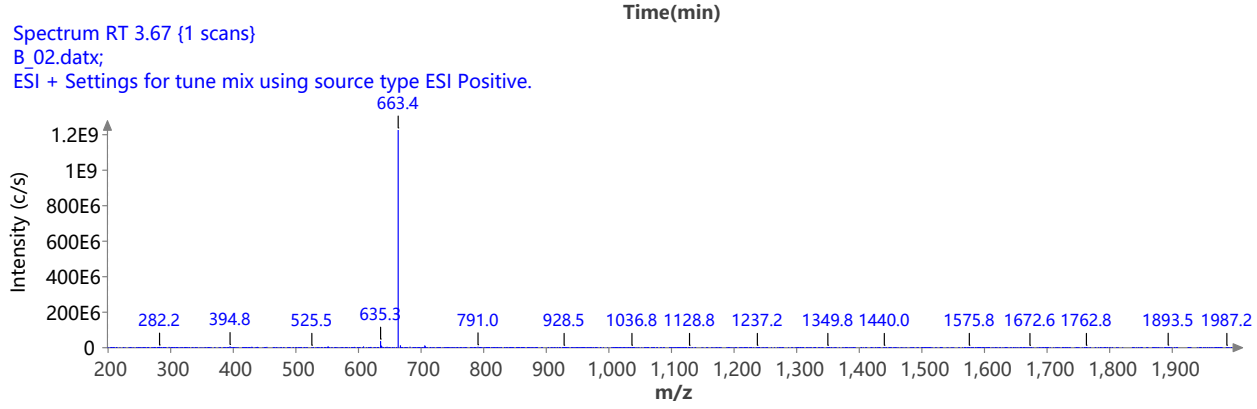
LCMS of compound **1-Pye**[6]



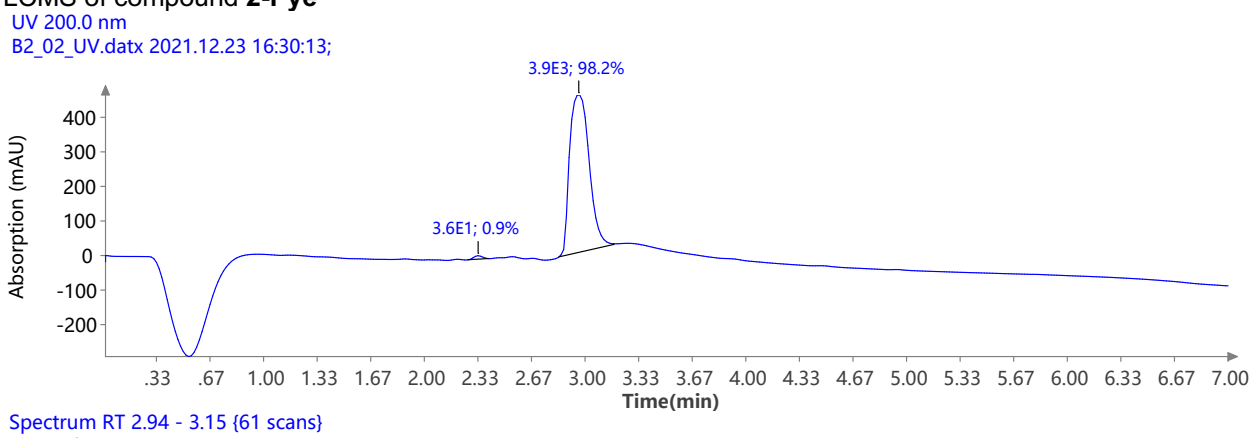


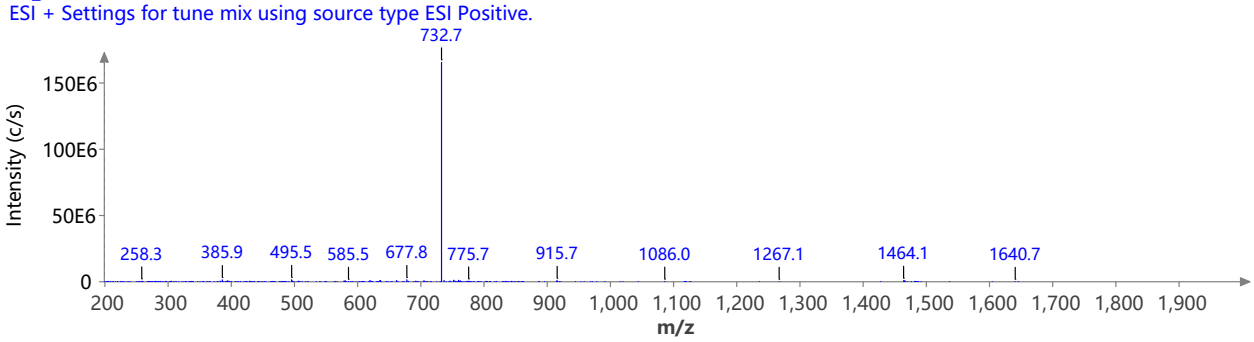
LCMS of compound **2**



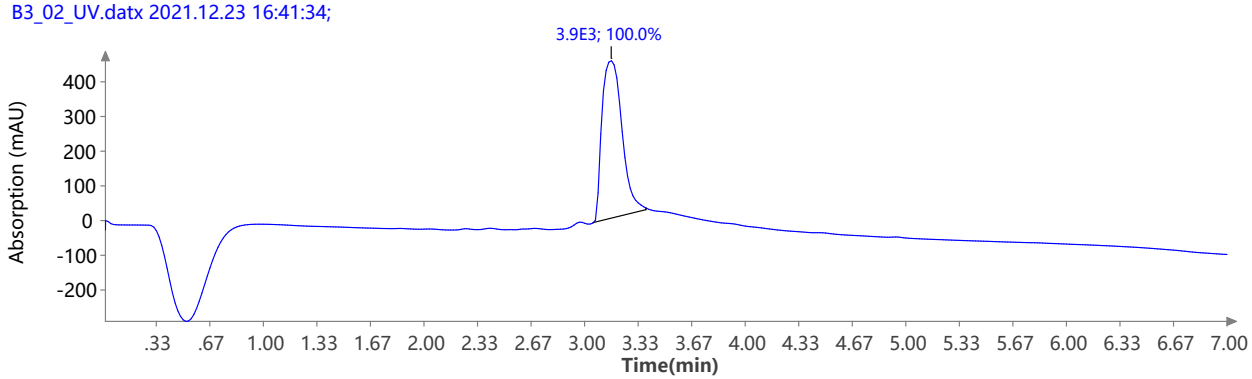


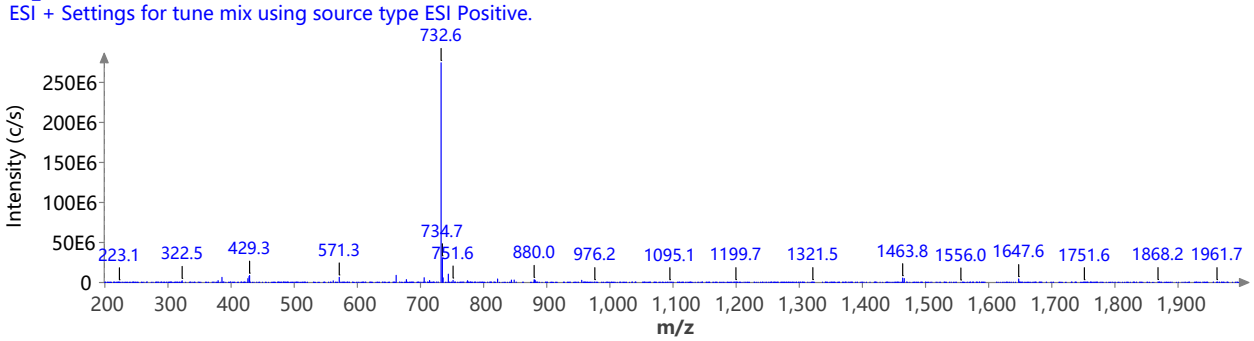
LCMS of compound **2-Pye^2**



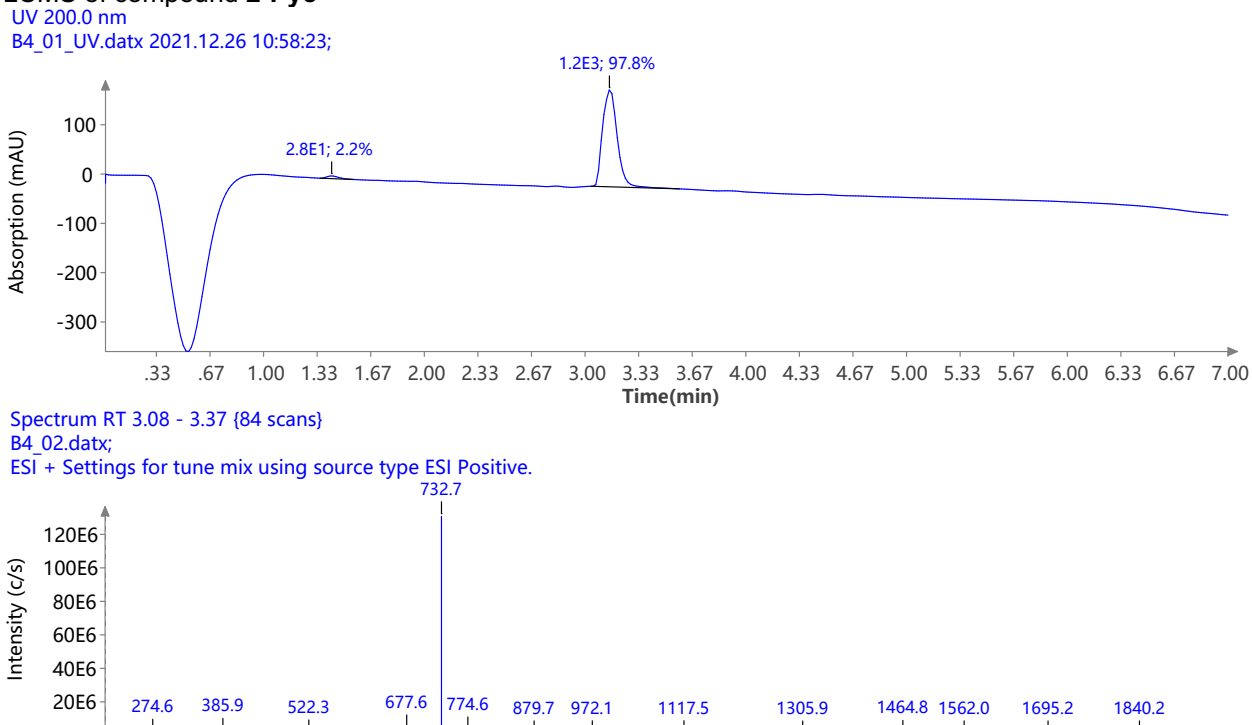


LCMS of compound **2-Pye^3**

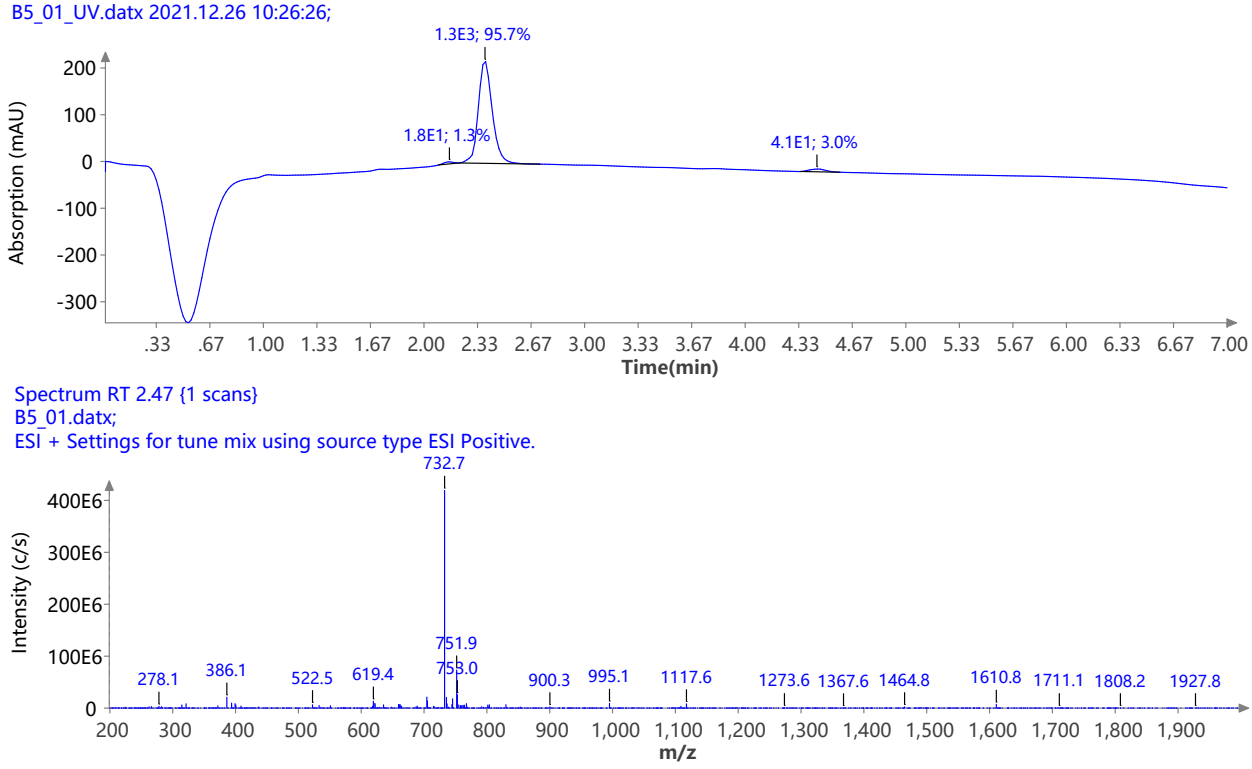




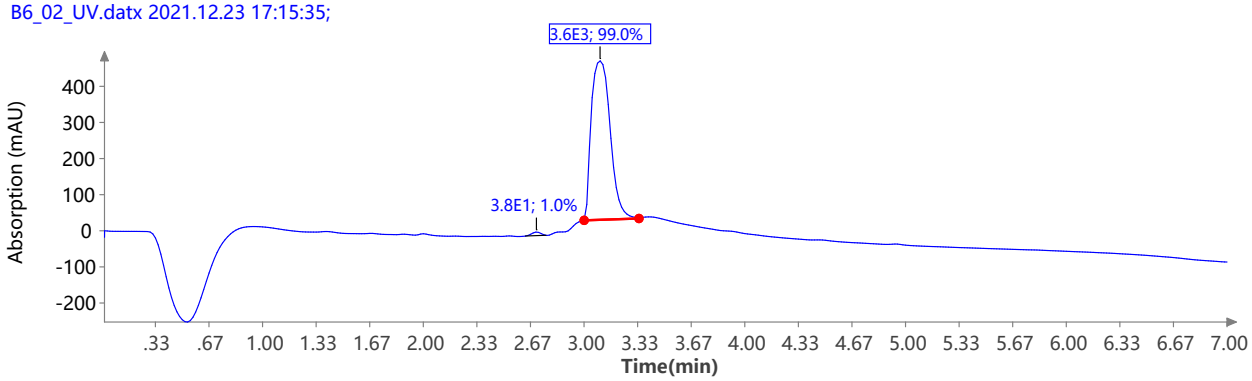
LCMS of compound **2-Pye**[4]

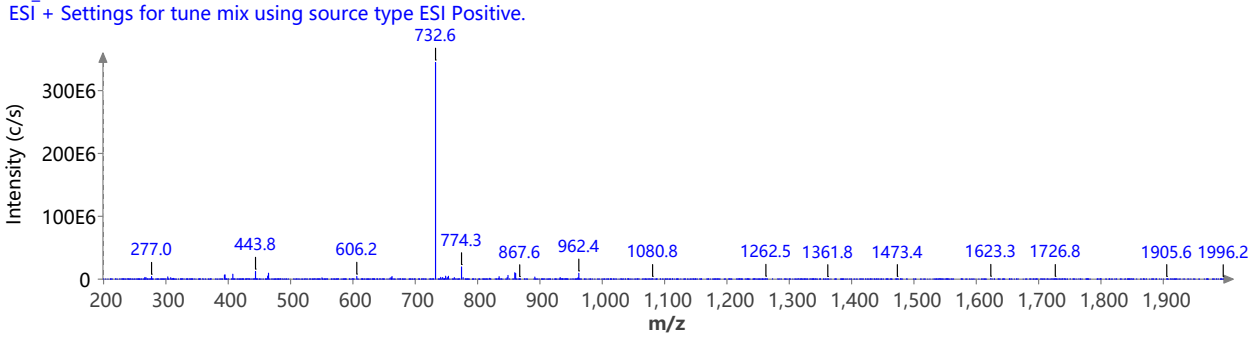


LCMS of compound **2-Pye**[5]

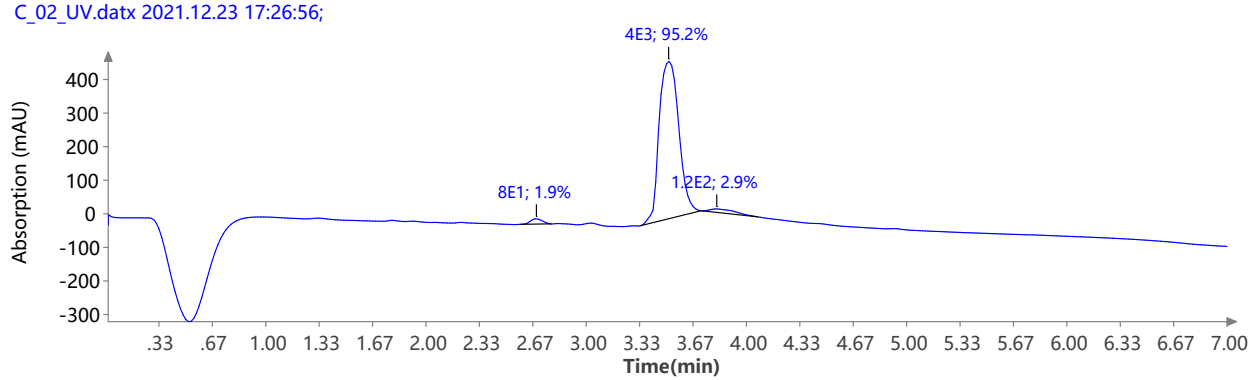


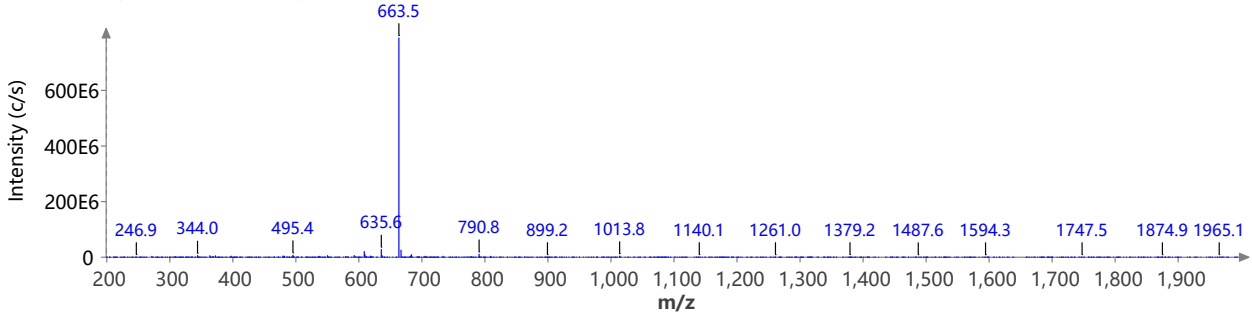
LCMS of compound **2-Pye**[6]



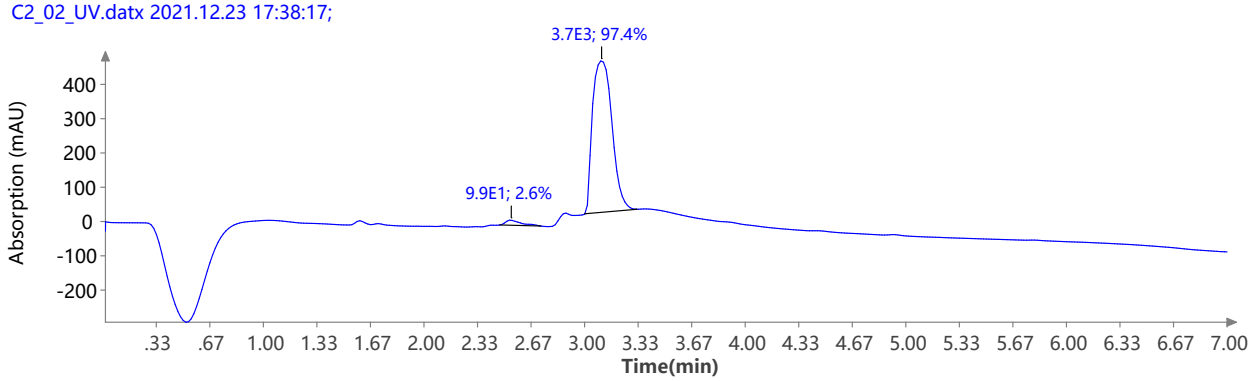


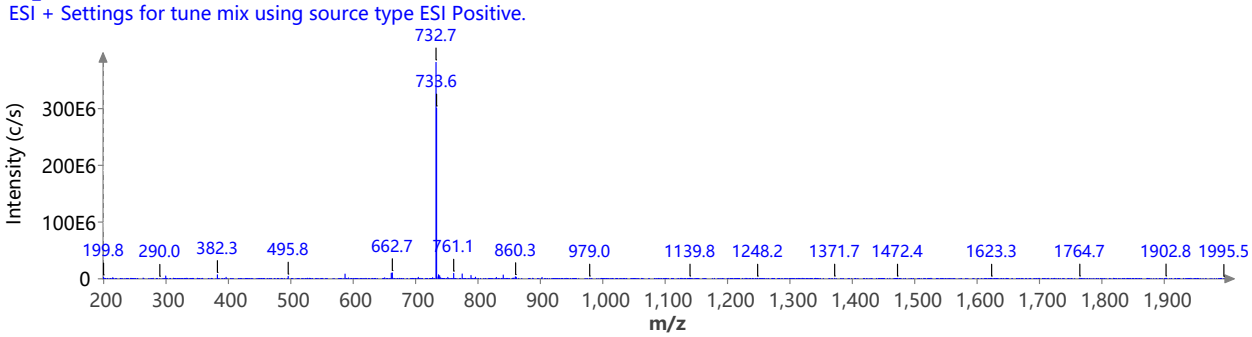
LCMS of compound **3**



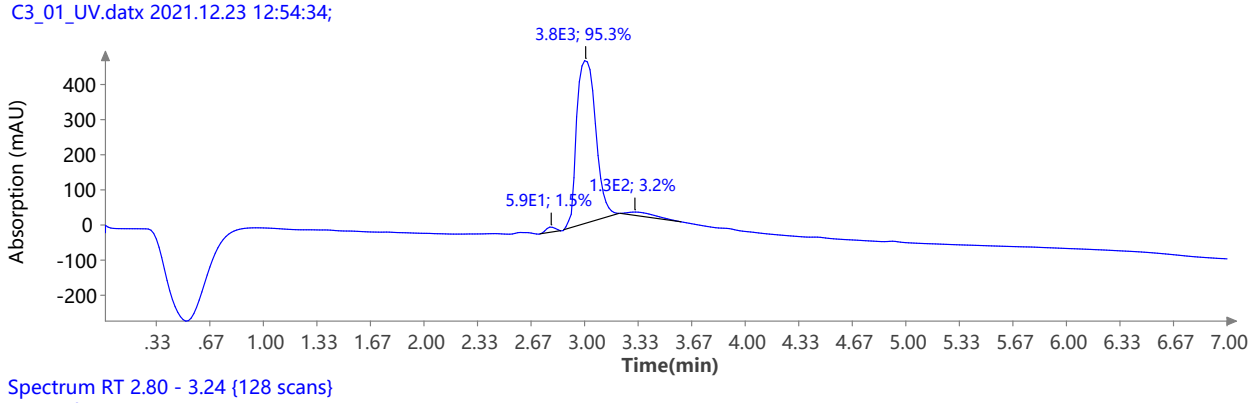


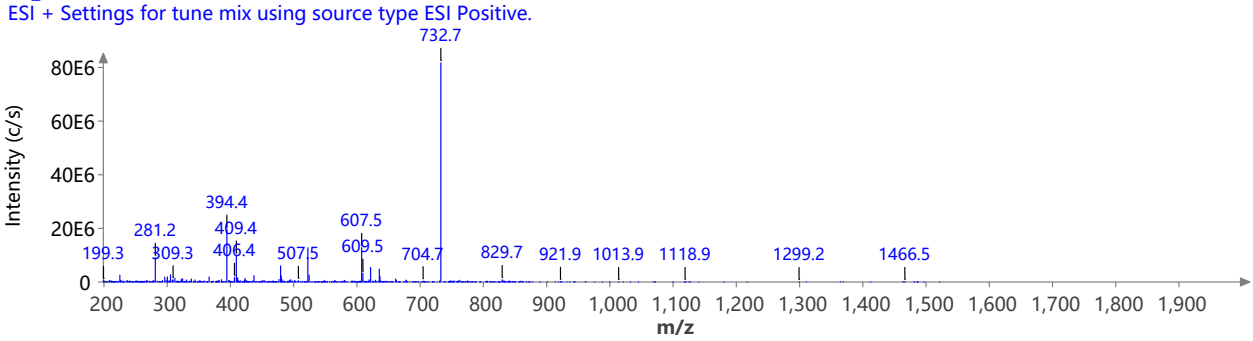
LCMS of compound **3-Pye^2**



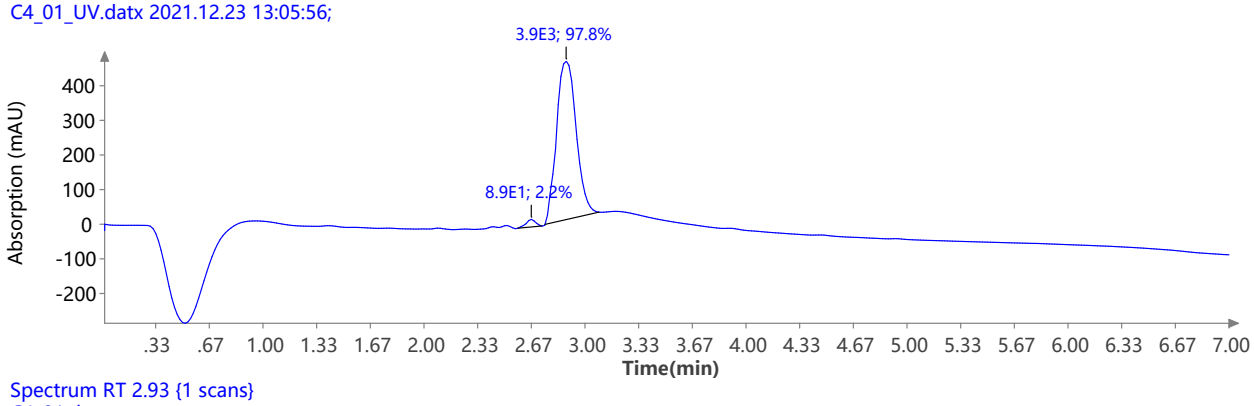


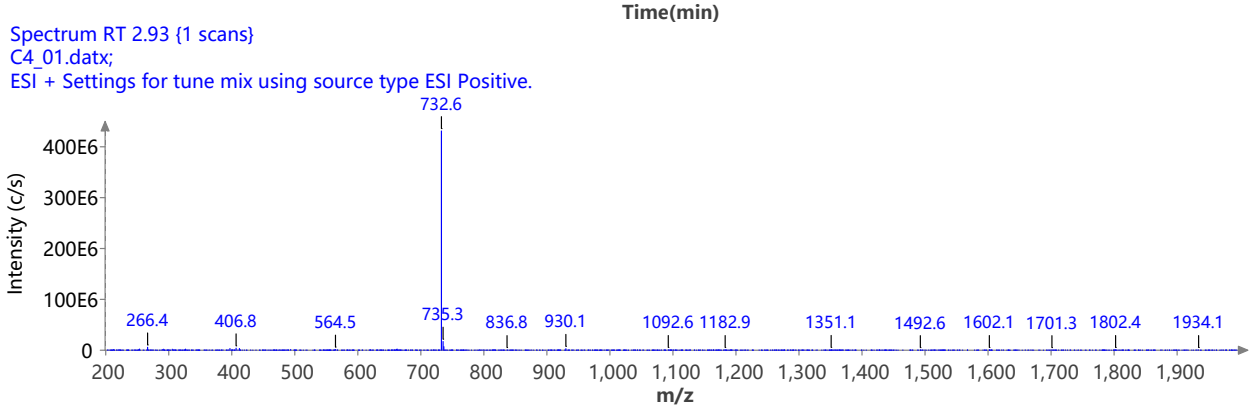
LCMS of compound **3-Pye^3**



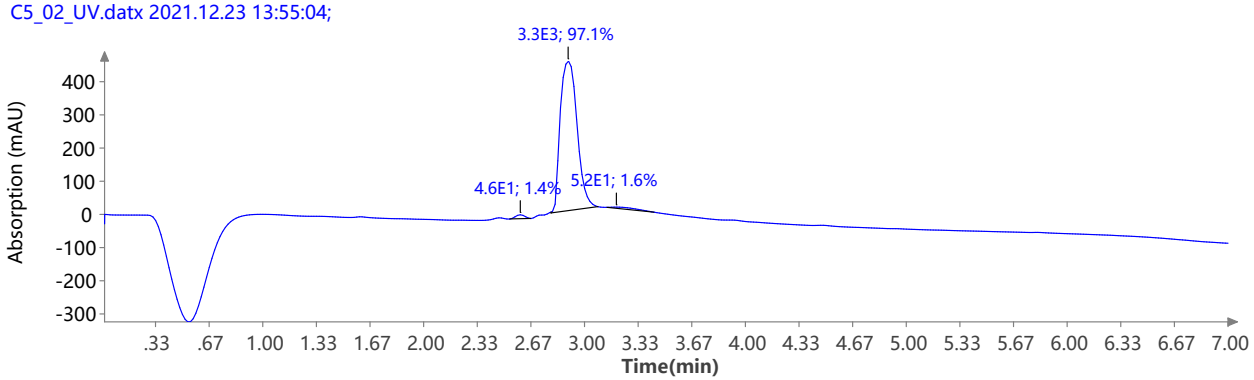


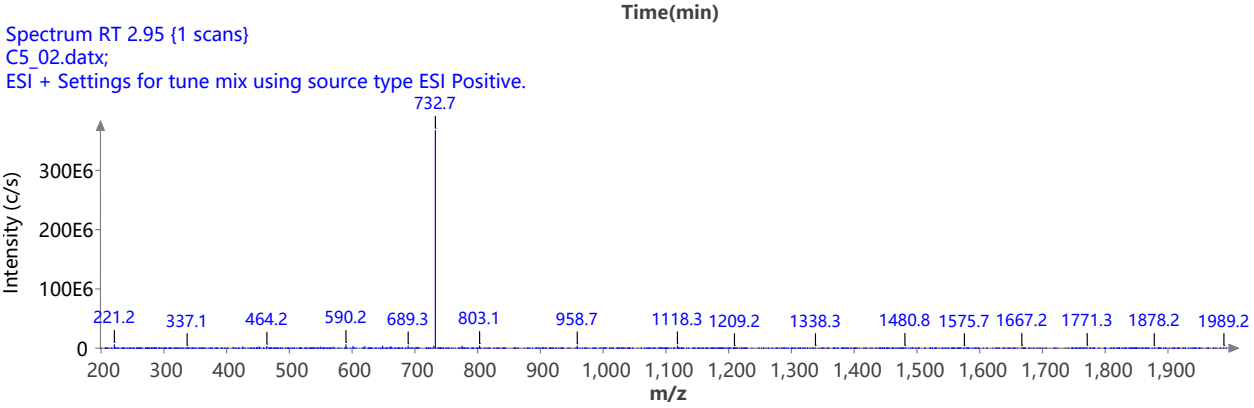
LCMS of compound **3-Pye[4]**



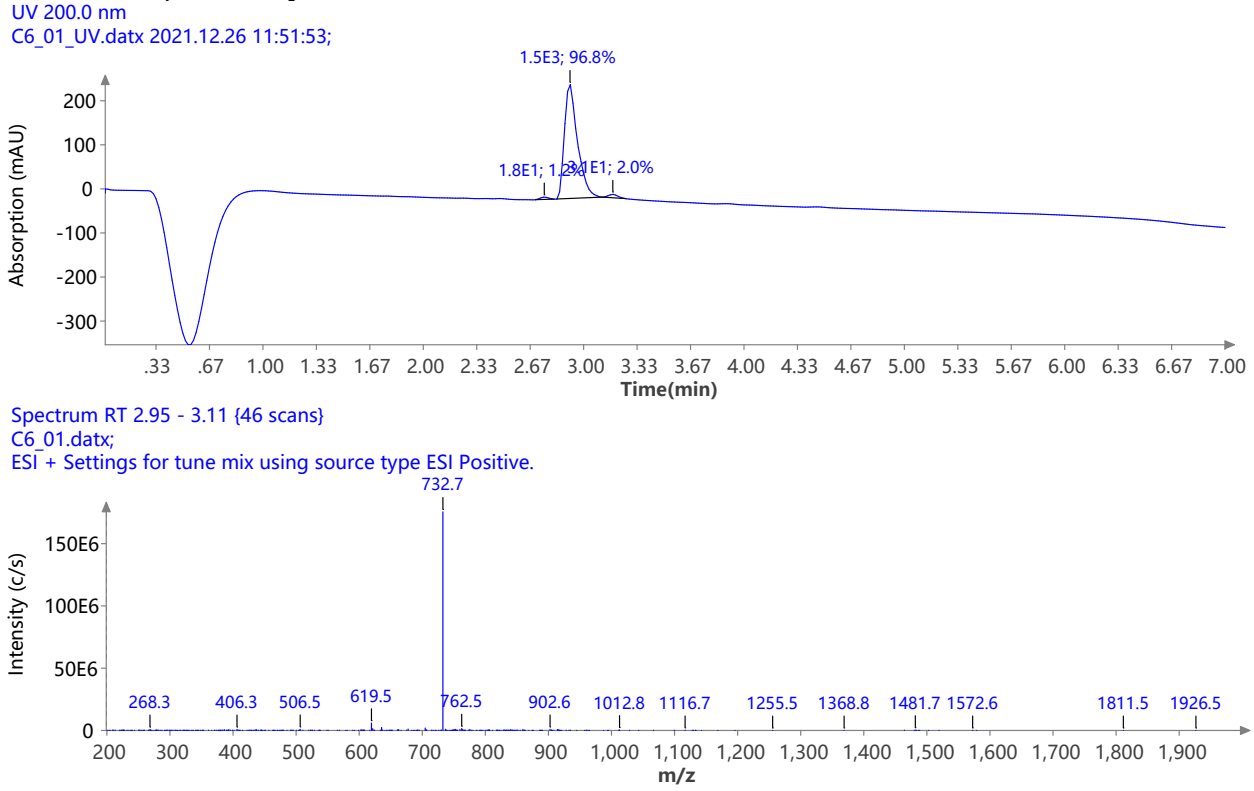


LCMS of compound **3-Pye[5]**

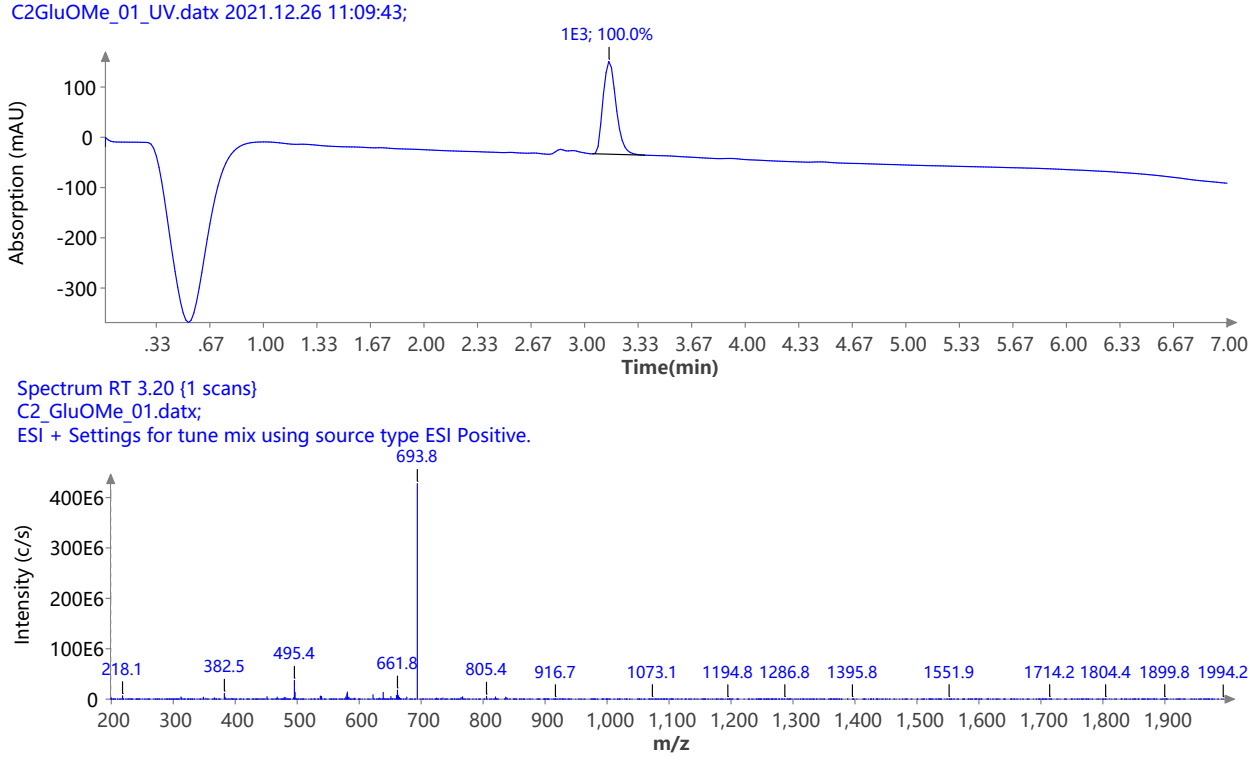




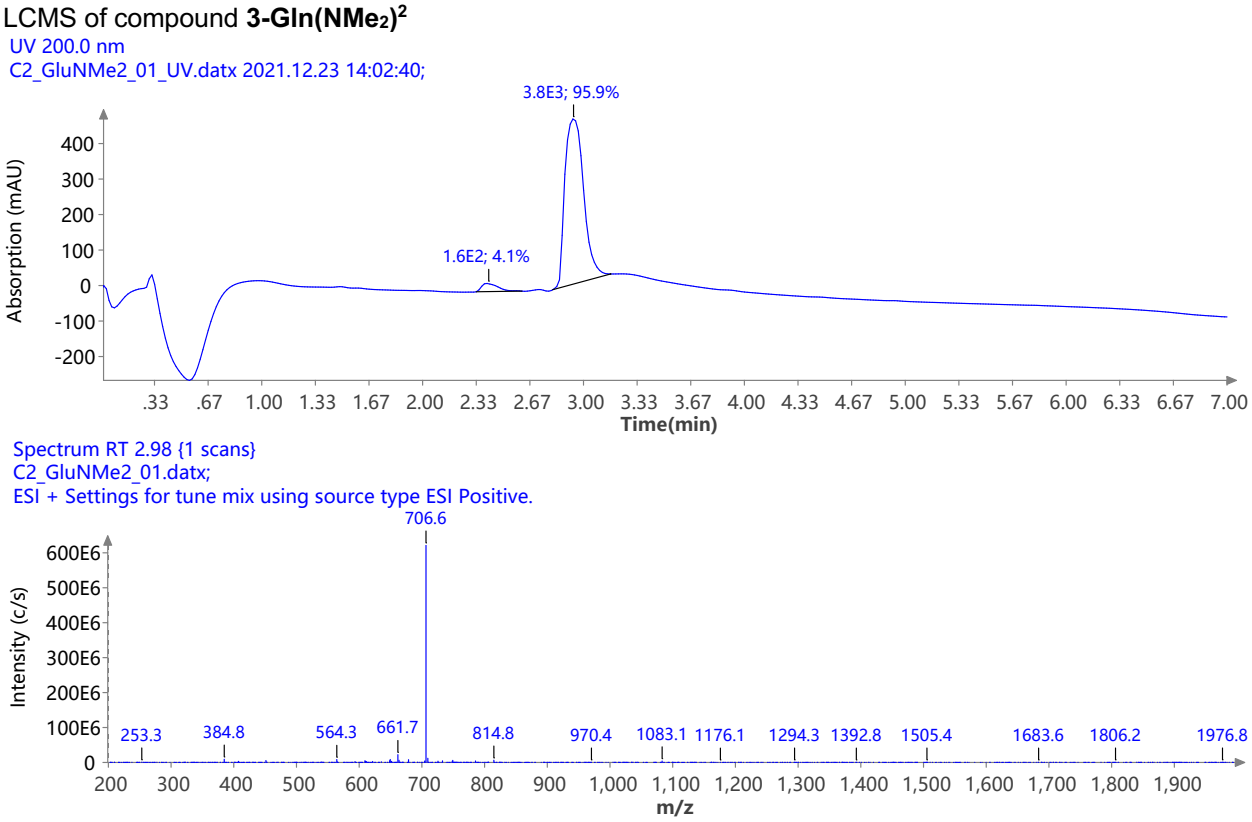
LCMS of compound **3-Pye**[6]



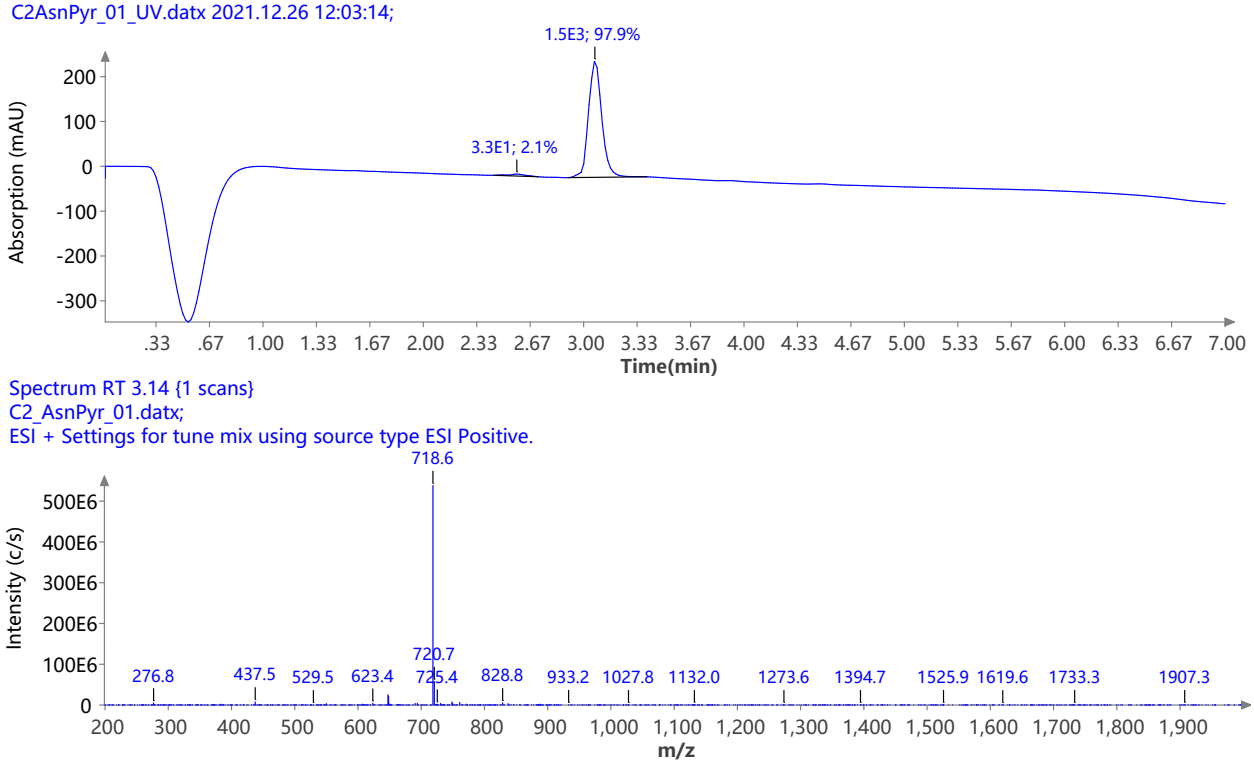
LCMS of compound **3-Glu(OMe)**[2]



LCMS of compound **3-Gln(NMe$_2$)^2**



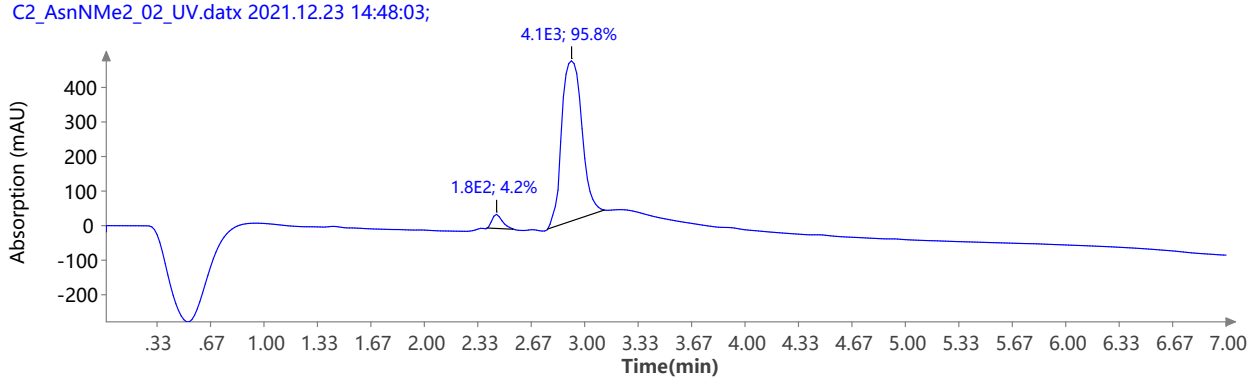
LCMS of compound **3-Asn(Pyr)^2**



LCMS of compound **3-Asn(NMe$_2$)^2**

UV 200.0 nm

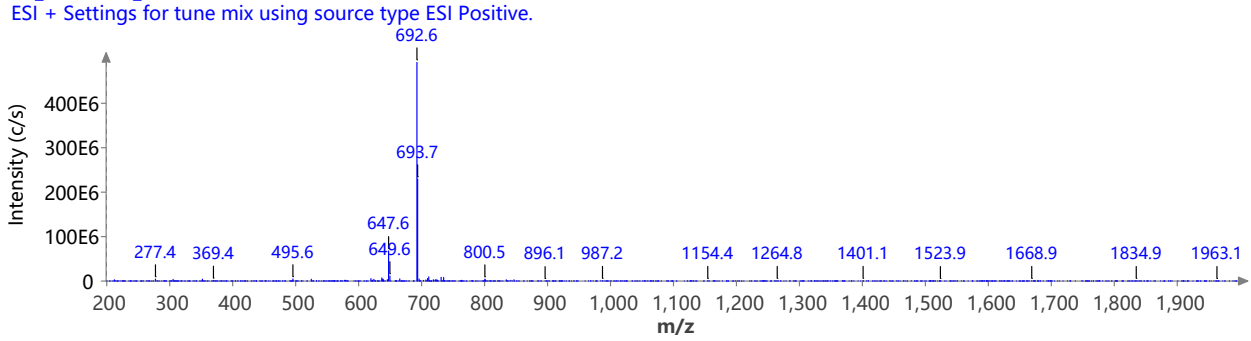
C2_AsnNMe2_02_UV.datx 2021.12.23 14:48:03;



Spectrum RT 2.98 {1 scans}

C2_AsnNMe2_02.datx;

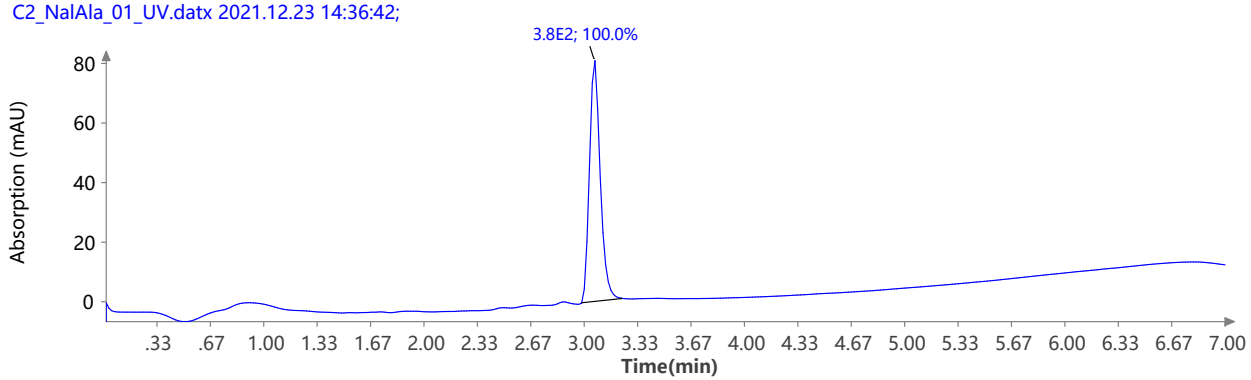
ESI + Settings for tune mix using source type ESI Positive.



LCMS of compound **3-Pye^2(Ala^{3}Nal^5)**

UV 254.0 nm

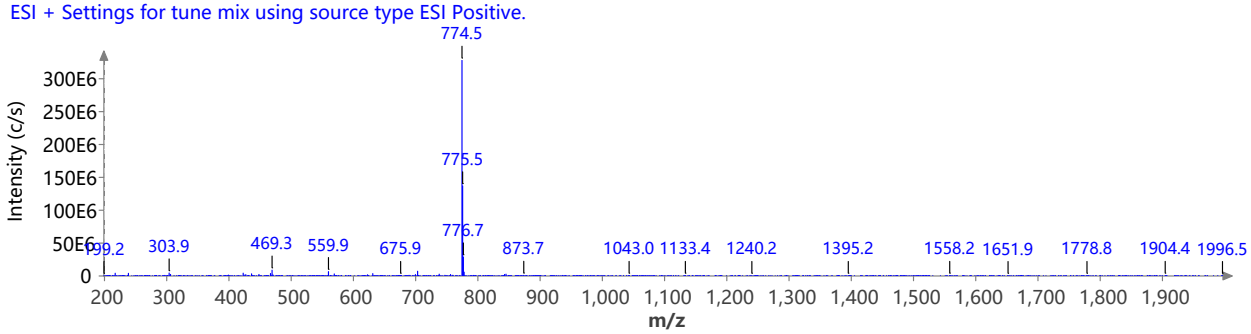
C2_NalAla_01_UV.datx 2021.12.23 14:36:42;



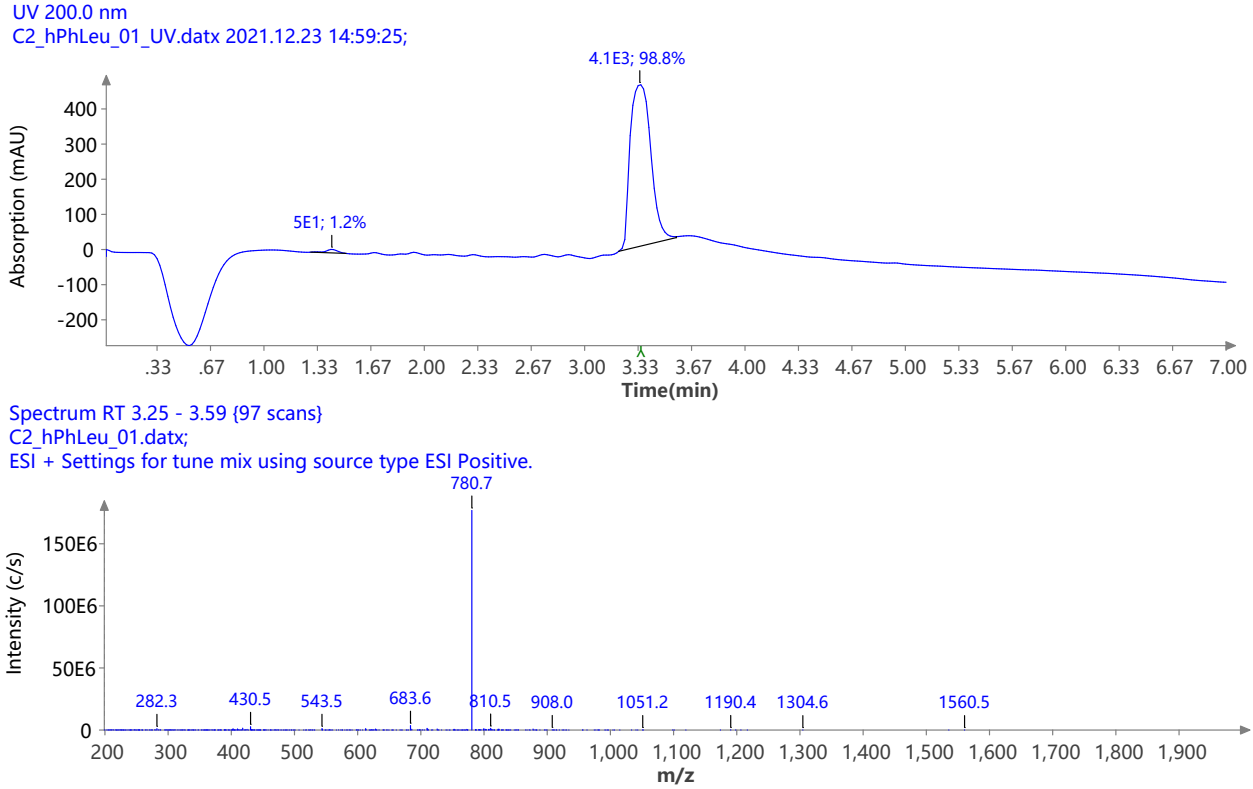
Spectrum RT 3.13 {1 scans}

C2_NalAla_01.datx;

ESI + Settings for tune mix using source type ESI Positive.



LCMS of compound **3-Pye^{2}(HPhe^{5})**



References

(1) Naylor, M. R.; Ly, A. M.; Handford, M. J.; Ramos, D. P.; Pye, C. R.; Furukawa, A.; Klein, V. G.; Noland, R. P.; Edmondson, Q.; Turmon, A. C.; Hewitt, W. M.; Schwochert, J.; Townsend, C. E.; Kelly, C. N.; Blanco, M. J.; Lokey, R. S. Lipophilic Permeability Efficiency Reconciles the Opposing Roles of Lipophilicity in Membrane Permeability and Aqueous Solubility. *J. Med. Chem.* **2018,** *61*, 11169-11182.

(2) Esposito, G.; Pastore, A. An alternative method for distance evaluation from NOESY spectra. *J. Magn. Reson. (1969)* **1988,** *76*, 331-336.

(3) Hu, H.; Krishnamurthy, K. Revisiting the initial rate approximation in kinetic NOE measurements. *J. Magn. Reson.* **2006,** *182*, 173-177.

(4) Butts, C. P.; Jones, C. R.; Towers, E. C.; Flynn, J. L.; Appleby, L.; Barron, N. J. Interproton distance determinations by NOE – surprising accuracy and precision in a rigid organic molecule. *Org. Biomol. Chem.* **2011,** *9*, 177-184.

(5) Begnini, F.; Poongavanam, V.; Atilaw, Y.; Erdelyi, M.; Schiesser, S.; Kihlberg, J. Cell Permeability of Isomeric Macrocycles: Predictions and NMR Studies. *ACS Med. Chem. Lett.* **2021,** *12*, 983-990.

(6) Rigaku Oxford Diffraction, CrysAlisPro software system, Rigaku Corporation, Wroclaw, Poland. **2020**.

(7) Sheldrick, G. SHELXT - Integrated space-group and crystal-structure determination. *Acta Crystallogr. A* **2015,** *71*, 3-8.

(8) Sheldrick, G. Crystal structure refinement with SHELXL. *Acta Crystallogr. C* **2015,** *71*, 3-8.

(9) Müller, P. Practical suggestions for better crystal structures. *Crystallogr. Rev.* **2009,** *15*, 57-83.