

# Global Analysis of Peptide Cyclization Efficiency

Amit Thakkar<sup>§</sup>, Thi Ba Trinh<sup>§</sup>, and Dehua Pei\*

## Supporting Information

**Table S1.** Sequences of fast-cyclizing peptides selected from libraries I-IV.

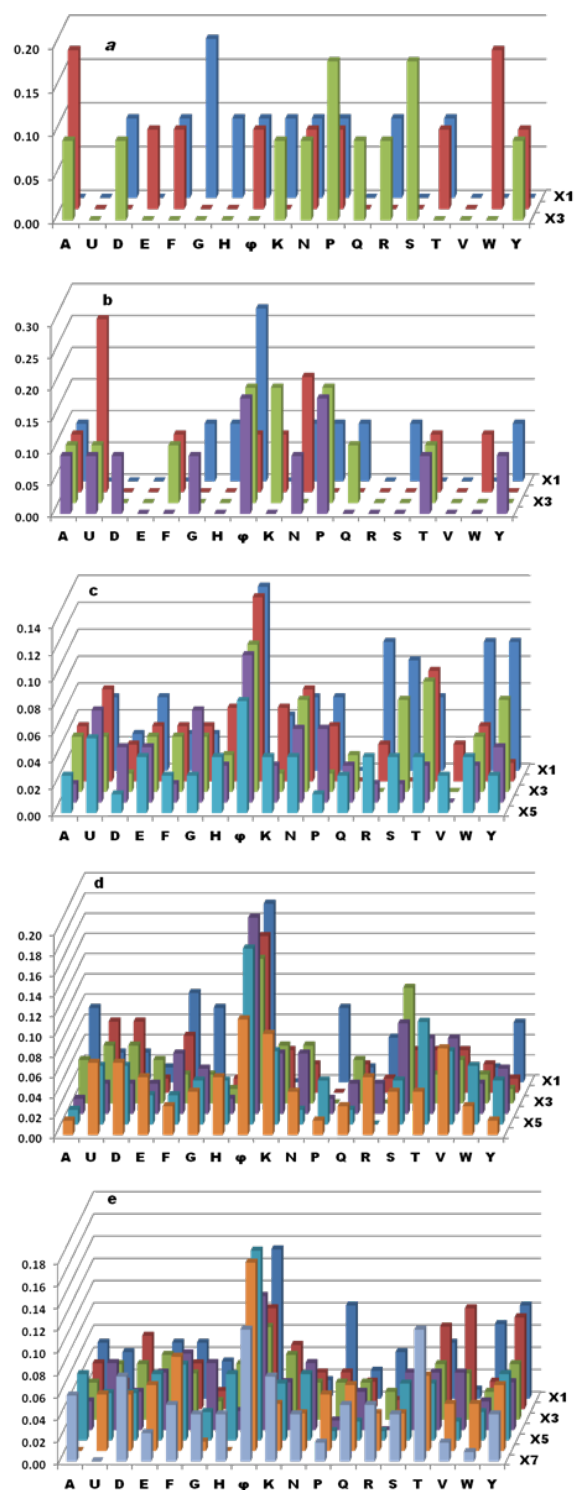
| Library I | Library II | Library III |        | Library IV |         |
|-----------|------------|-------------|--------|------------|---------|
| SAKE      | GTUYE      | SEHUWE      | XXXXAE | φTRAφEE    | HDEWXXE |
| RADE      | UUUQE      | WDφURE      | XXXXSE | PSUAFFE    | WEGNXXE |
| NFTE      | φAUSE      | NYWEYE      | XXXUPE | SφφEWQE    | KFφSφAE |
| SφGE      | NFFPE      | FRFGφE      | XXXXWE | GφSGTφE    | XXXXφGE |
| PNPE      | TKKAE      | ADYφYE      | WφSUFE | UTKSFSE    | φWSφRYE |
| APGE      | YφNNE      | NUφφDE      | KNφGUE | RPDTTHE    | WYφSXXE |
| QTφE      | AKNφE      | φENφDE      | QHAφNE | XUTKFSE    | φVφφKGE |
| PWNE      | DQTHE      | GGDNHE      | XXXTKE | FKYXXXX    | VWφFXXE |
| DWFE      | PφWφE      | EUKPWE      | HGGKUE | XXXXXφE    | QUφφφDE |
| YYHE      | φPAGE      | UAPφRE      | TPFNφE | XXKVXXE    | SφFXXPE |
| KERE      | PPφφE      | RTWφφE      | XXSKPE | RGEWSFE    | YVφUφWE |
|           |            | NφEGφE      | TPφXXE | TQWφQφE    | UHAXXXE |
|           |            | SPFFEE      | XXQTφE | HSφKAφE    | VNφKφUE |
|           |            | XXXXYE      | RφSTKE | EUFSSRE    | TWYTHφE |
|           |            | EUUHφE      | XXXDWE | HGφNXXE    | AKHXXXX |
|           |            | φHNNYE      | NNEASE | SφNVφHE    | NTTφQRE |
|           |            | φGYφWE      | φWYWHE | KYUVVφE    | EφTQVφE |
|           |            | XXXXPE      | XXXNφE | GφDφTYE    | φYKXXXX |
|           |            | XUTTEE      | EQTRRE | KEKφUφE    | UVSφφφE |
|           |            | WPGEFE      | FXXXXE | EDNNWYE    | RVVQφAE |
|           |            | TNWKTE      | XXXXRE | FPPQSAE    | VPφXXXX |
|           |            | XXXHRE      | HYNVTE | VφSFUGE    | KSSφUPE |
|           |            | XXXXφE      | YDYPSE | KGNSEGE    | XXXXXFE |
|           |            | PGTWφE      | XXXXKE | XXXXXAE    | KKGEUAE |
|           |            | XXXφGE      | DGTAYE | VφVSAGE    | φφVXXXX |
|           |            | UFYKEE      | HQTAHE | UVQYφUE    | UDSφURE |
|           |            | AWφWNE      | XXNTGE | XXXXKPE    | φHφENFE |
|           |            | SφEHWE      | RKφHSE | KKNNXXE    | RφVSXXE |
|           |            | XXXXPE      | YTHRφE | VUQGXXE    | TKTANPE |
|           |            | GKφEYE      | UφSPYE | QφYHFFE    | NWφSDYE |
|           |            | XXXφNE      | UφSYUE | HTWDYφE    | NHRXXXX |
|           |            | XXAφNE      | φφGTRE | φFDKVφE    | GφFUKFE |
|           |            | XXUVTE      | φUAXXE | DAVUφφE    | φTGXXXX |
|           |            | VφQUSE      | XXXXRE | ETFUYAE    | DTTXXPE |
|           |            | VSTFSE      | KENFWE |            |         |
|           |            | KYUNTE      | XXXXUE |            |         |
|           |            | QNφDEE      |        |            |         |

φ = Ile, Leu and Nle; U = Abu; X = amino acid whose identity could not be unambiguously determined.

**Table S2.** Sequences of fast cyclizing peptides selected from library V.

| Library V  |            |           |           |
|------------|------------|-----------|-----------|
| TRALHNSE   | XXXXXXNPE  | TWKVPAφE  | KDYDXXXXE |
| GDUFYXXE   | SEQφYGE    | XXAVVVφE  | RφYUUGFE  |
| MFDENARE   | φTKAφPWE   | φPφφNPWE  | RφEφKVAE  |
| TSDUMYTE   | PEφSXXXE   | RφHφφDφE  | KYSXXXXE  |
| DFEQHXXE   | FFFHWDEE   | TNφFφDYE  | FQFφQHTE  |
| NUETUYPE   | WYUφAHφE   | XWYGEKPE  | DEHXXXXE  |
| APFYXXPE   | QRφφTSEE   | DDHWSYYE  | TWNKQPFE  |
| ALHTFQSE   | φYAVDVUE   | φφTVUVφE  | YUXXXXXE  |
| TFHFEWHE   | KFUPYφWE   | TEAEKφTE  | KYENFφEE  |
| IFIMDQYE   | ETKATVRE   | SYNNKTφE  | φVφFXXXXE |
| TQISGXE    | XXXXXXXXGE | HRφEUKφE  | φKPFYTRE  |
| SWKSEYQE   | GNGKHKEE   | φTDφDφGE  | DRφSDKPE  |
| KILIRTTE   | XXXXXXXXPE | TKTGTTWE  | FφφUSUφE  |
| HEMIHYPE   | QNSFFAE    | HFGNHYAE  | KφφFRTPE  |
| MESGTSQE   | XXXXXGTE   | NPφUφφTE  | HφXXXXXE  |
| QISIUDIE   | VQYNYVWE   | YφFφEKQE  | DFGVRWφE  |
| ADYTDNφE   | VTφTYVYE   | QφKUETEE  | PφSXXXXE  |
| TKWFEPFE   | φEφφXXXXE  | GφAFETWE  | KφφGXXXXE |
| KTφφDφYE   | φφφYQFYE   | XFTWVVφE  | RSSDUTHE  |
| TPNXXXXE   | ATWSVDAE   | φφφXXXXE  | APφXXXXE  |
| φNφQKφφE   | RQFφGDUE   | SVUKQARE  | XXAφφφFE  |
| AφDYITFE   | DφRKHDWE   | FYDYNYYE  | φUHφKVφE  |
| XXXXXXKUE  | TTYEφφUE   | XXXXXSANE | YφGUWVPE  |
| YVTKANφE   | EDNNFUKE   | TφNVP UAE | EVTAFY YE |
| AFQGAFFE   | TφTYARNE   | KSFTφVφE  | DUQXXXXE  |
| YXXXXXXXXE | NWHXXXXE   | QKYEWEHE  | NGφXXXXE  |
| FXXXXXXXXE | XXXXXXXXEE | SVEKTFYE  | GSXXXXXE  |
| XφFNφφφE   | XXXXXXXXPE | GDAHVKAE  | NKVφNEUE  |
| DQNQSWVE   | RTWGKFEE   | FPUDPEDE  | φQEUPDWE  |
| XXXTKφGE   | DYFWVARE   | φUφNTQKE  | HQVEφYYE  |
|            |            |           | QUEDXXXXE |

φ = Ile, Leu and Nle; U = Abu; X = amino acid whose identity could not be unambiguously determined.



**Figure S1.** Histograms showing the amino acid composition/sequence of fast cyclizing peptides selected from libraries I (a) to V (e). The y axis value represents the frequency (maximum value 1.0) at which each amino acid (x axis) was selected at a given random position ( $X^1$  to  $X^7$  on the z axis).  $\phi$  = Ile, Leu and Nle. U = Abu.

**Table S3.** Sequences of poorly cyclizing peptides selected from libraries I-IV.

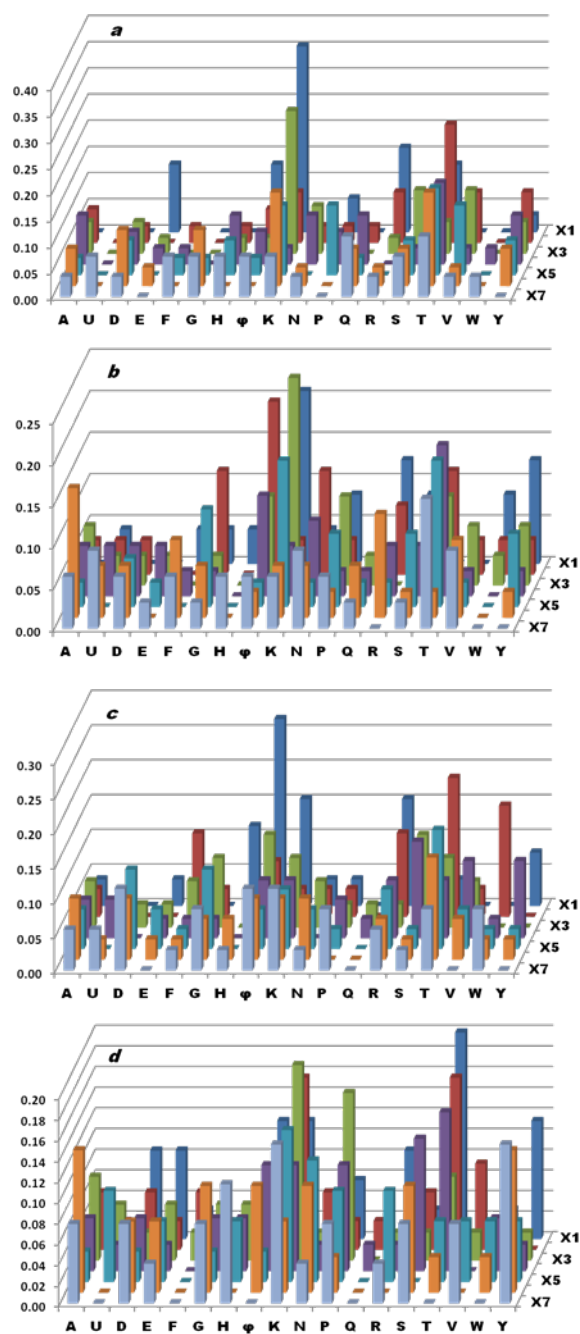
| Library I | Library II | Library III |        | Library IV |         |
|-----------|------------|-------------|--------|------------|---------|
| φKUE      | AQVFE      | YDGFPE      | FφNAKE | ASHFRKE    | APKNWRE |
| GNFE      | FFKGE      | HPWHPE      | ENFFRE | SKKFRHE    | YRSXXXE |
| GUKE      | φPARE      | HKEφUE      | HGKTKE | KKQHKHE    | XXTWKHE |
| AHKE      | KVTPE      | HHKKRE      | VHKφKE | HRKQVFE    | YKKφNRE |
| TRKE      | KTDNE      | GKKKFE      | KHKTKE | TWφKVRE    | AHFKWRE |
| TTKE      | TKYAE      | FYKRNE      | PARGPE | WRKKHSE    | KYYXXXE |
| XTKE      | VKRKE      | φARRRE      | GSRRFE | VHPKFKE    | TPNXXXE |
| TTKE      |            | KHYRRE      | TKNFKE | TYYRVFE    | TKGRXXE |
| GRNE      |            | RWWRFE      | RQYKRE | YYKRYQE    | XXXKGYE |
| KRPE      |            | TKTRRE      | HFKTKE | KUURRFE    | TYAXXXE |
| SKRE      |            | TWYRRE      | KNNRRE | TKYXXXE    | ARAFRφE |
| XKRE      |            | φYRWFE      | PHKTKE | GAKHEAE    | KVTYRRE |
| XKRE      |            | TKNRKE      | TTKRKE | YKUHUEE    | PYYEφRE |
| KQRE      |            | YTKRUE      | KYXXXE | YFGRKYE    | YRKGRφE |
| GRSE      |            | TKYRRE      | TFKTKE | UNXXXXE    | RKFTYRE |
| URSE      |            | RHTTKE      | RφYRKE | UHAREFE    | TWTFHRE |
| KPTE      |            | QRUKRE      | φYNRRE | TYKXXXE    | FRTRXXE |
| TKYE      |            | HKKTKE      | KRFRNE | GYYXXXE    | YφRNKKE |
|           |            | RYGRKE      | VHRKKE | YKFφRFE    | RYKAFKE |
|           |            | KYRRPE      | XXXXXE | KTKYXXE    | GHKKRYE |
|           |            | TFRKGE      | φRKNφE | ARFKQYE    | RYSARNE |
|           |            | TKXXNE      | RRGWYE | TYFXXXE    | FφRRTKE |
|           |            | VRKYRE      | XXXXXE | YYSPKRE    | YYSRRKE |
|           |            | TKKTPE      | PSHRRE | RKFHRφE    | YφRNEKE |
|           |            | ERYRKE      | YKYTRE | TKFφWKE    | φKKWRNE |
|           |            | KKHYRE      | YGRRφE | AYRKSYE    | RQYWNRE |
|           |            | KYFKHE      | YRNYVE |            |         |
|           |            | φRNWVE      | WRKφQE |            |         |
|           |            | YFXXX E     | XGKRφE |            |         |
|           |            | RNDFYE      | VQSSRE |            |         |
|           |            | KYφHRE      | YRGRYE |            |         |
|           |            | GTKXXE      | WRYφKE |            |         |
|           |            | XXYRFE      | RVYYKE |            |         |
|           |            | RTVNFE      | WKFKKE |            |         |
|           |            | RHYFRE      | FTRRYE |            |         |
|           |            | YTNFNE      | RPWKRE |            |         |
|           |            | TKSRNE      | VYRRHE |            |         |
|           |            | KTHYRE      | RYGTRE |            |         |
|           |            | KYGGHE      | KYTWRE |            |         |

φ = Ile, Leu and Nle; U = Abu; X = amino acid whose identity could not be unambiguously determined.

**Table S4.** Sequences of poorly cyclizing peptides from library V

|                        |                        |                       |                             |
|------------------------|------------------------|-----------------------|-----------------------------|
| TWQNAVAE               | KNK $\phi$ FH $\phi$ E | TKUUEWNE              | $\phi$ UHGKGRE              |
| NVKWKWUE               | PPGPGK $\phi$ E        | GAKFAKNE              | TUTKUUSE                    |
| SSGNP $\phi$ UE        | PTKRDK $\phi$ E        | Y $\phi$ SFVTPE       | KFTWRSSE                    |
| YKYFTKUE               | A $\phi$ HYHR $\phi$ E | RFUTNKPE              | KYGKHYSE                    |
| WRQTYKUE               | PPUNK $\phi$ $\phi$ E  | TUYTKRPE              | $\phi$ KSG $\phi$ $\phi$ SE |
| PYLKNEUE               | UTK $\phi$ KH $\phi$ E | FREERTPE              | IQTTQKSE                    |
| HDNSKPEE               | AHPKYGKE               | SY $\phi$ RKKPE       | YRYRDHTE                    |
| ETHDGIEE               | YRWNSYKE               | RSFQKHPE              | KGGHRPTE                    |
| XPNKRKKE               | TNYTK $\phi$ KE        | TUTKVKPE              | UHRNRPTE                    |
| UQKFQTFE               | PNGKPTKE               | QEKVVPPE              | WTKTKNTE                    |
| RTTQKHFE               | HRNUKTKE               | KY $\phi$ NVKPE       | NSPD $\phi$ PTE             |
| FKK $\phi$ VWFE        | $\phi$ FNHHGKE         | DEKKEWPE              | KTTKKSTE                    |
| HAD $\phi$ SRFE        | Q $\phi$ FHTTKE        | FKKDLLQE              | FPYKR $\phi$ TE             |
| RRFHQUFE               | TFQNPKE                | TSRFQDQE              | Y $\phi$ KH $\phi$ RTE      |
| RIKKKKFE               | HTK $\phi$ KGKE        | MATPTIQE              | RWQRGAW                     |
| HRUTWWGE               | TKKVSYKE               | TWTN $\phi$ KRE       | FNNWPKWE                    |
| RFNKYHGE               | ETHKPTKE               | RVPYYHRE              | YSRP $\phi$ KWE             |
| NQHK $\phi$ RGE        | HEKTYUKE               | YYEHRYRE              | NV $\phi$ DTPWE             |
| RFKYHKGE               | RRUTGRKE               | KNKGYRE               | PNNRRQWE                    |
| K $\phi$ YKQHGE        | AKYAFRKE               | TVKNGHRE              | KANAADWE                    |
| KKNWKVHE               | TWRQHVKE               | APFNVKRE              | VFKGGQWE                    |
| YTKRWKHE               | YYGKKTKE               | HVUWNVRE              | GHGKTRYE                    |
| YTWYRHHE               | GYNKNHKE               | HTAKGKRE              | N $\phi$ FDKYE              |
| UKYA $\phi$ RHE        | $\phi$ KRDKTKE         | UTPYNVRE              | YKYWRTYE                    |
| YNKF $\phi$ HHE        | UQPHK $\phi$ KE        | GKYKNYRE              | NSTWSRYE                    |
| RHGKQRHE               | $\phi$ KYKTNKE         | UTKPFARE              | SYTKSKYE                    |
| GYQKUK $\phi$ E        | GD $\phi$ QNVKE        | SYRGNNRE              | PTGYKGYE                    |
| KHHKQR $\phi$ E        | RKUYNYKE               | ATUTK $\phi$ RE       | DENKTFYE                    |
| DKFERR $\phi$ E        | NASHDWKE               | UTTUKRRE              | PUHGHYE                     |
| QKT $\phi$ NV $\phi$ E | UPYKFTKE               | $\phi$ KSKEHRE        | WRK $\phi$ SVYE             |
| WKHH $\phi$ K $\phi$ E | ADGIRKKE               | $\phi$ NY $\phi$ PRRE | TTRKKQYE                    |
| FHSHRK $\phi$ E        | TAKAFUKE               | THGHKKRE              | PTNYKTYE                    |
| UTUTKV $\phi$ E        | UHRVKKKE               | YGHGKKRE              | SPANFKYE                    |
| GKRPRK $\phi$ E        | KYTSSTKE               | A $\phi$ KAKRRE       | QKKRUUYE                    |
| KEFYGF $\phi$ E        | T $\phi$ NKYNNE        | PPKYDNRE              | RNKTEWYE                    |
| RADE $\phi$ R $\phi$ E | KUTPYGNE               | YKGERRRE              | YKKTGVYE                    |
| KYSPNF $\phi$ E        | TYKTK $\phi$ NE        | YNVWTKRE              | $\phi$ HRAAKYE              |
| $\phi$ FKYSK $\phi$ E  | RYHTTVNE               | QYSYTDRE              |                             |
| RAGKRW $\phi$ E        | YRKREKNE               | GAFVYKRE              |                             |
| VDUEFK $\phi$ E        | AYHKKYNE               | N KAKFIRE             |                             |

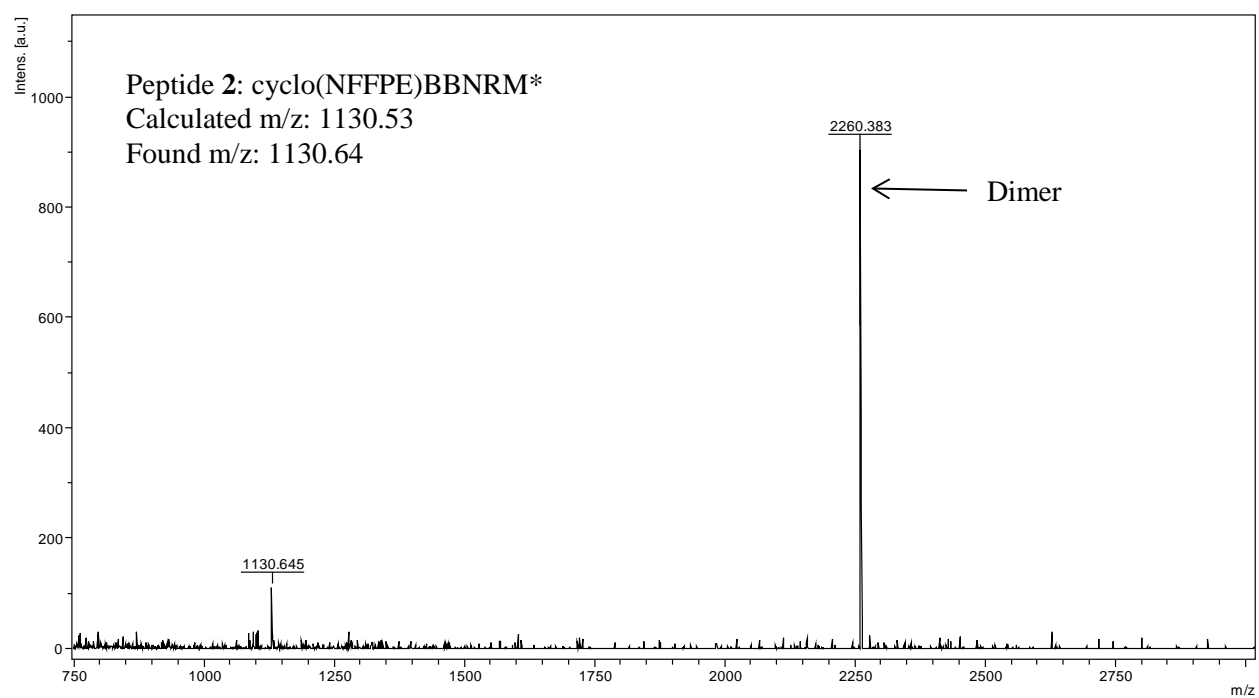
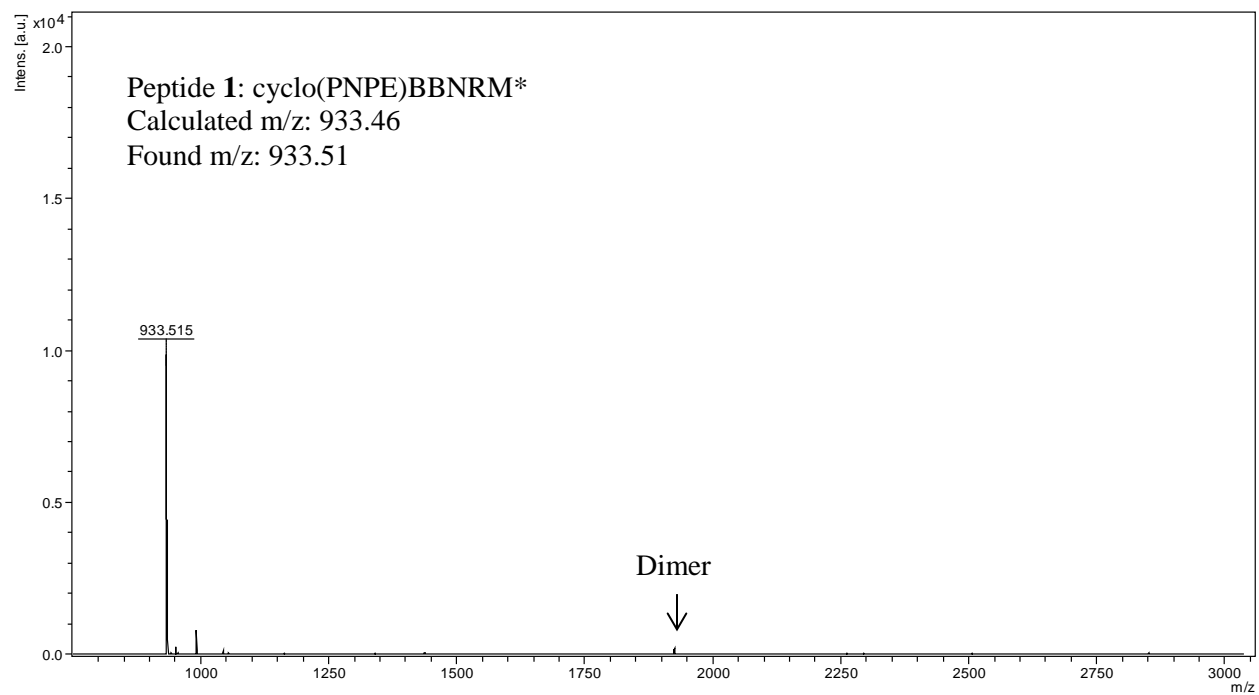
$\phi$  = Ile, Leu and Nle; U = Abu; X = amino acid whose identity could not be unambiguously determined.

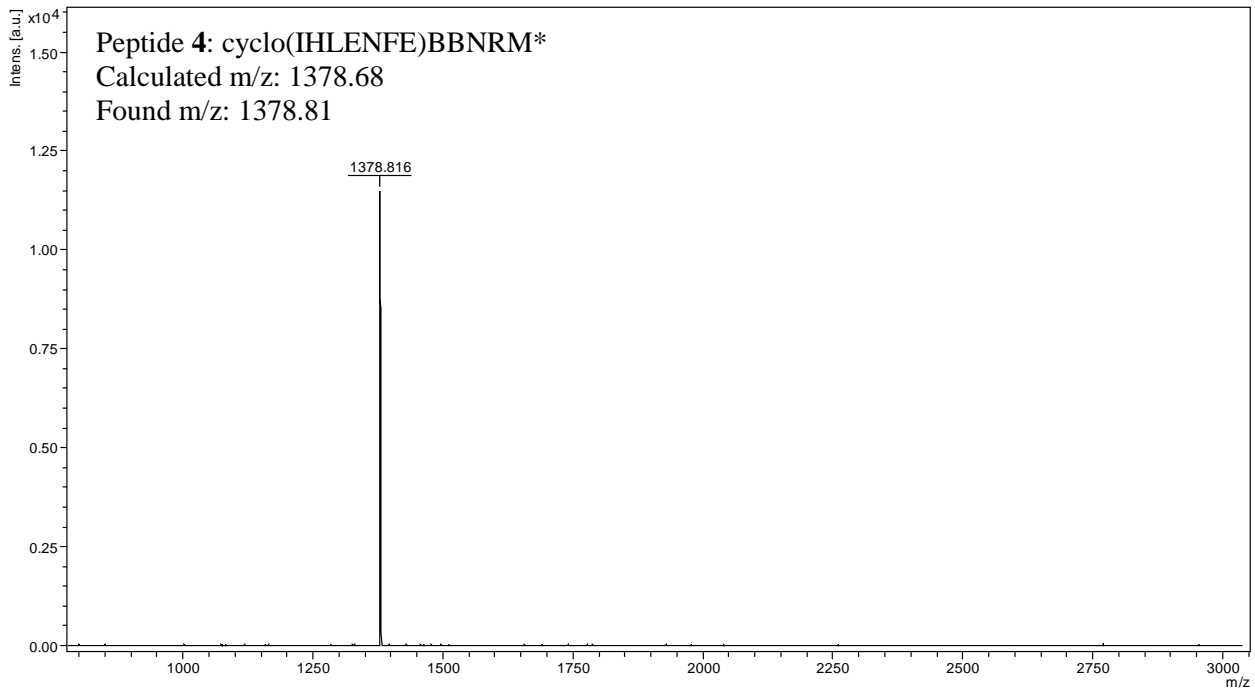
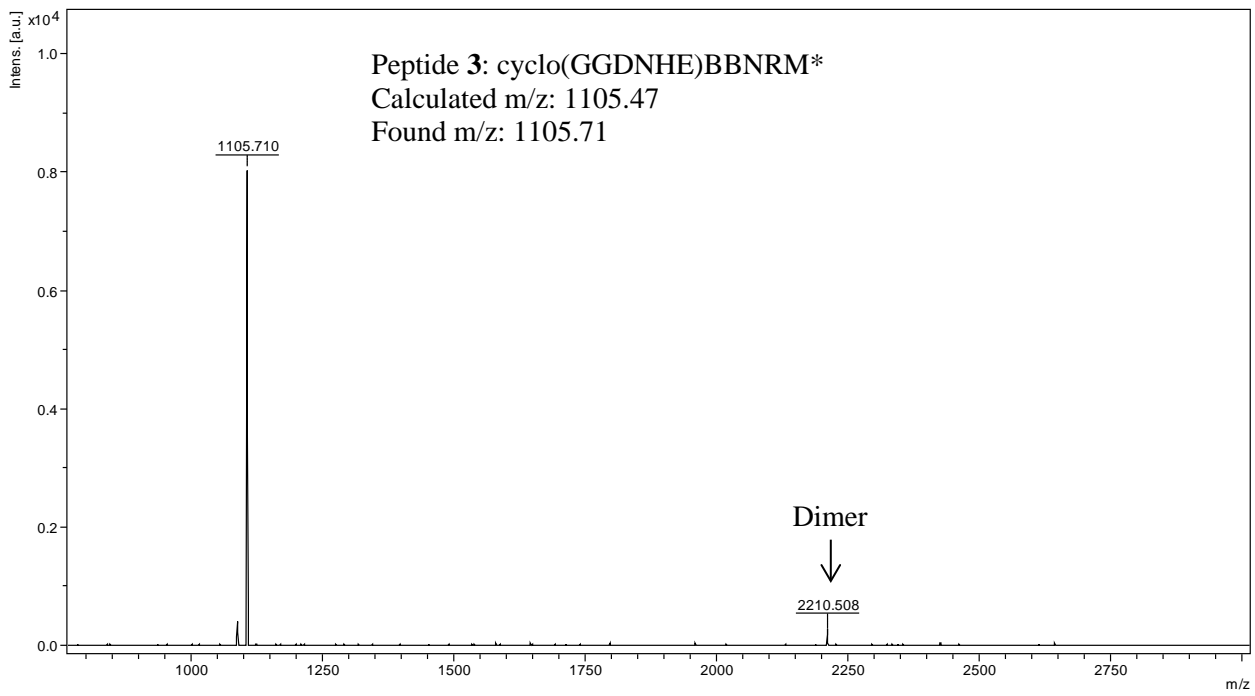


**Figure S2.** Histograms showing the amino acid composition/sequence of poorly cyclizing peptides selected from libraries VI (a) to IX (d). The y axis value represents the frequency (maximum value 1.0) at which each amino acid (x axis) was selected at a given random position (X<sup>1</sup> to X<sup>7</sup> on the z axis). φ = Ile, Leu and Nle. U = Abu.

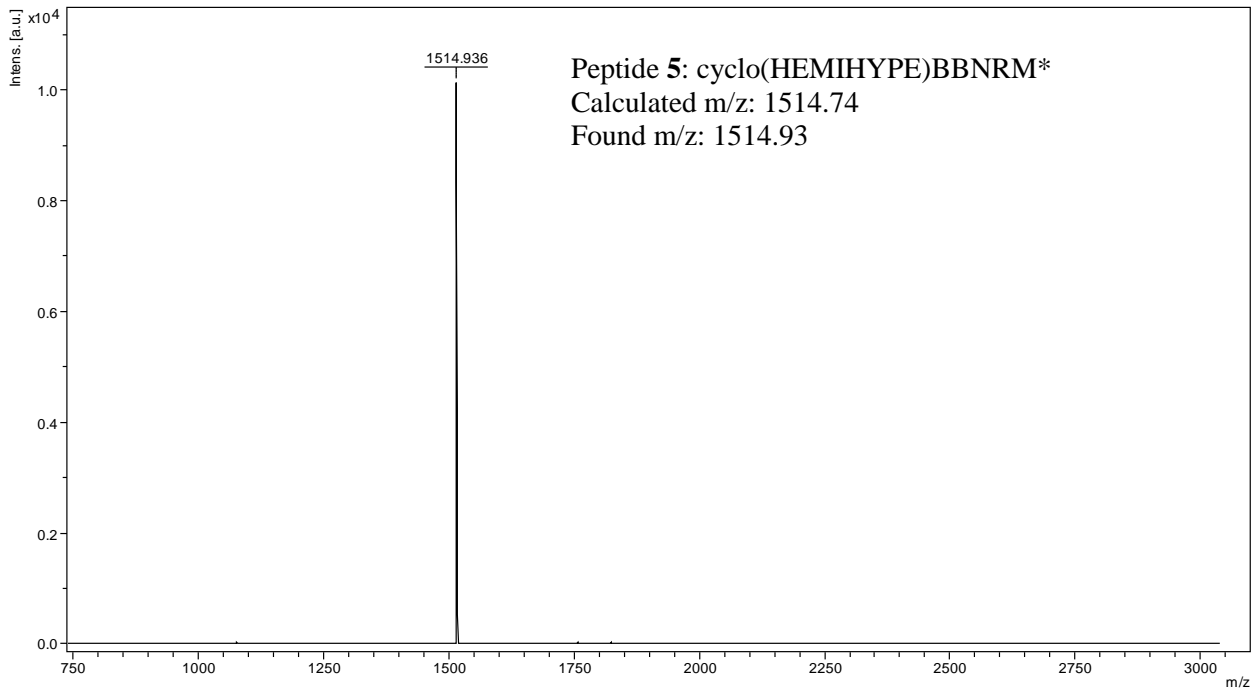
**Figure S3.** Representative MALDI-TOF MS spectra for some of the resynthesized fast- and poorly cyclizing peptides in Table 2. M\*, homoserine lactone; M, norleucine.

**Fast Cyclizing Peptides (1-5)**

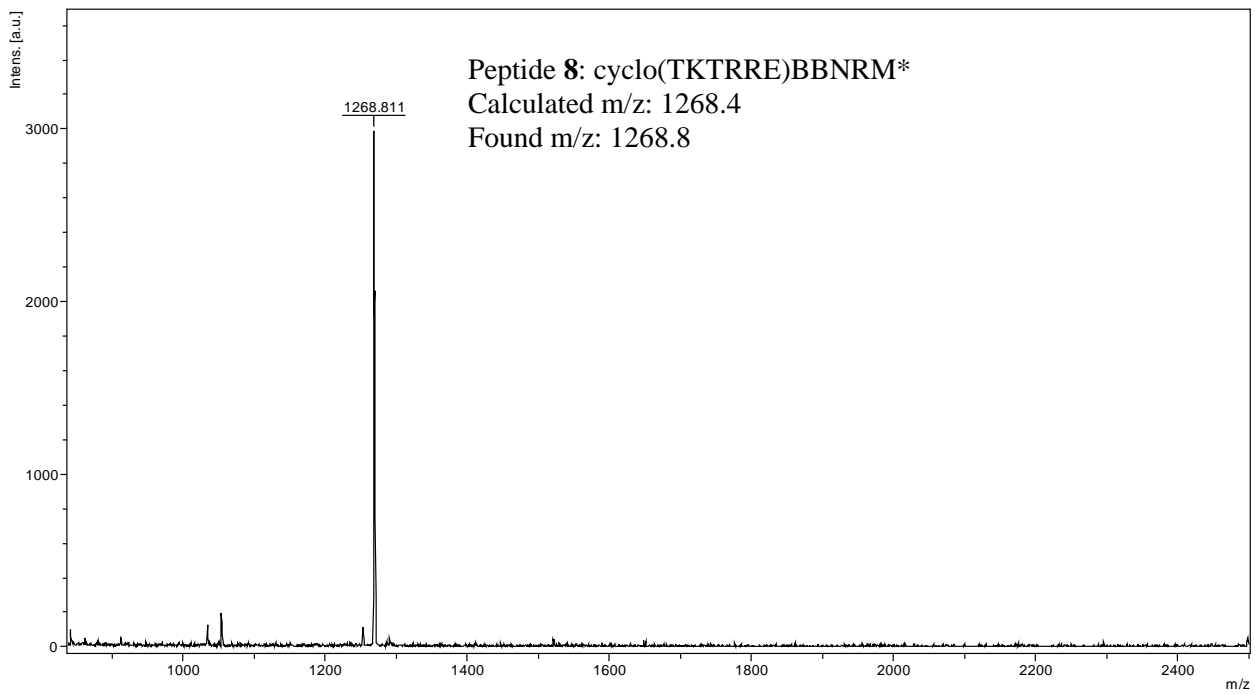


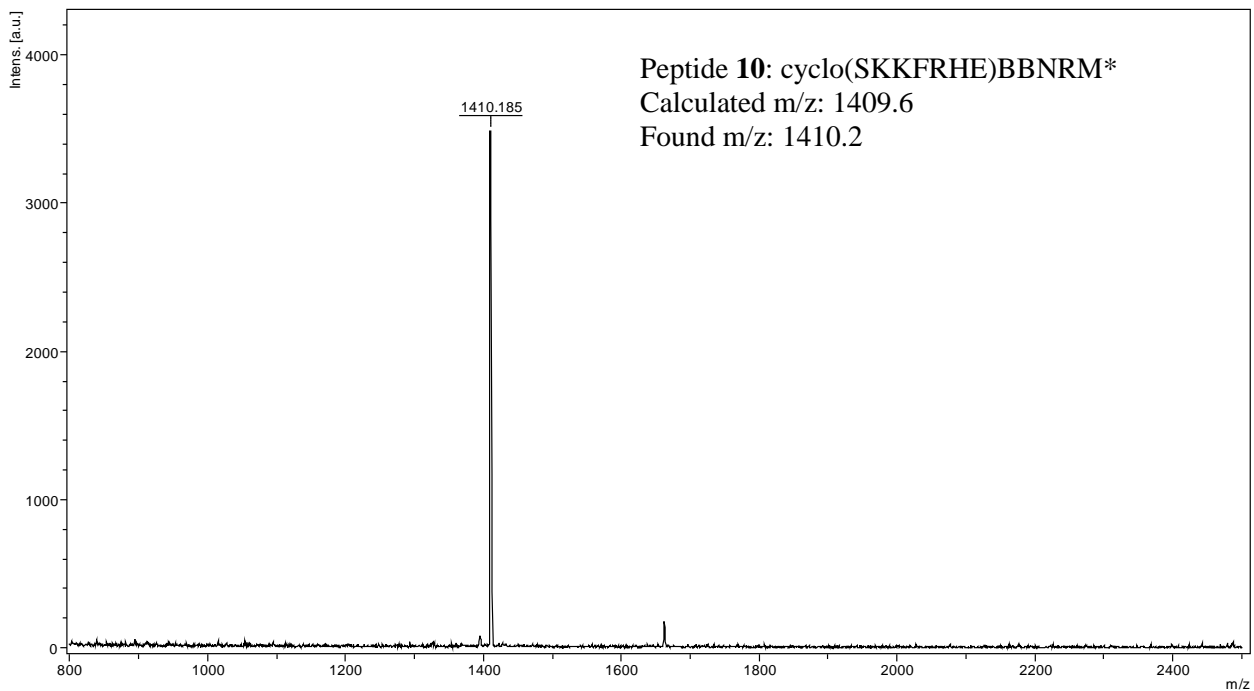






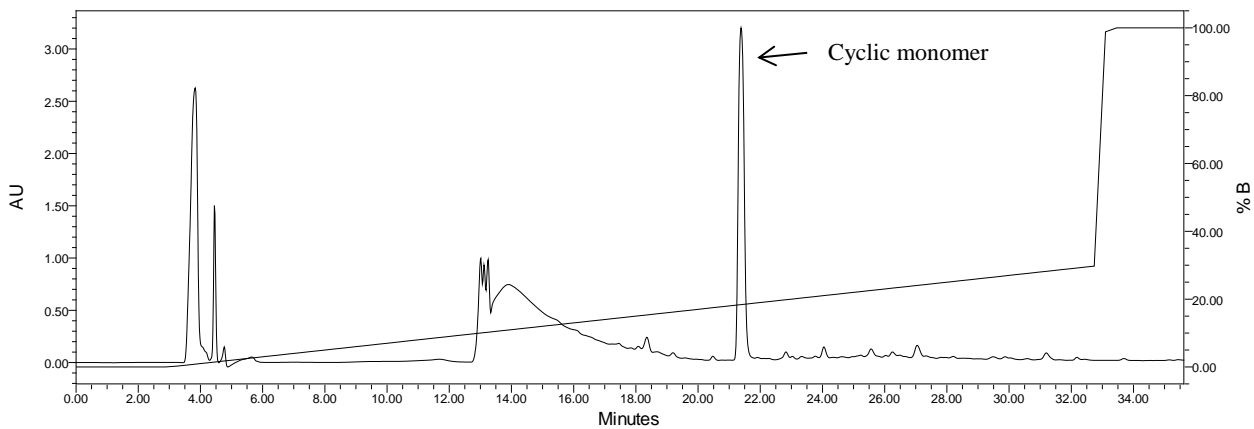
Representative Poorly Cyclizing Peptides (**8** and **10**)



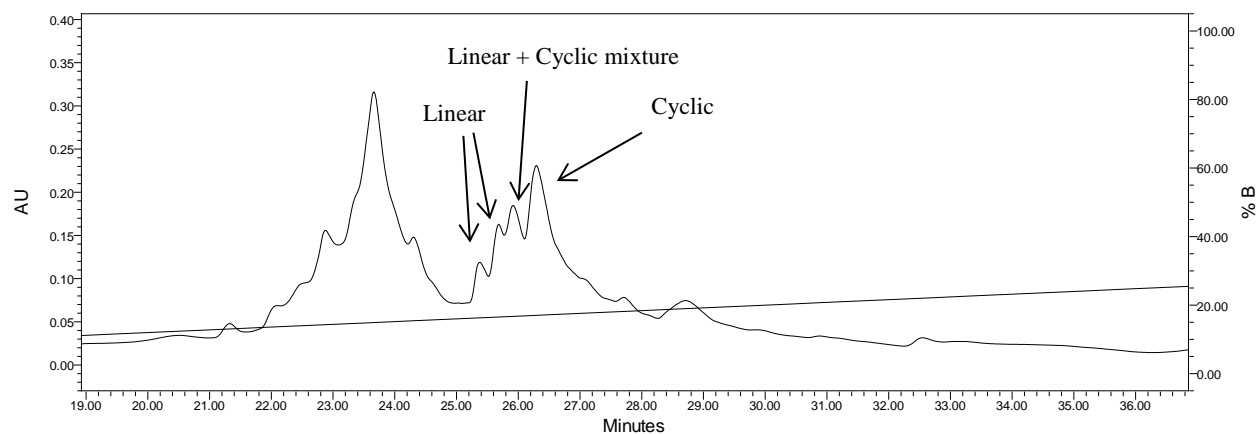


**Figure S4.** Representative HPLC chromatograms for some of the resynthesized poorly cyclizing peptides in Table 2. M\*, homoserine lactone; M, norleucine.

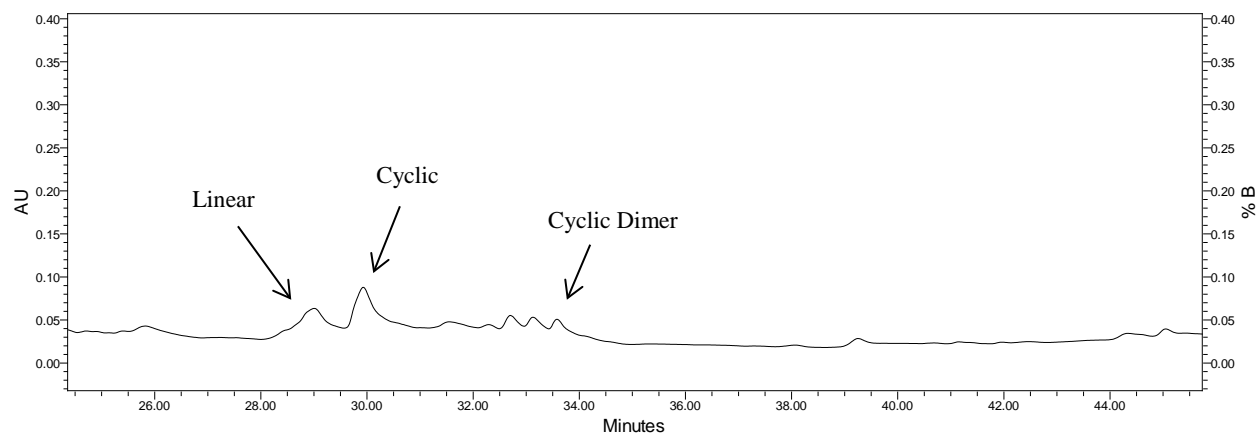
Peptide **7**: cyclo(TKYAE)BBNRM\*



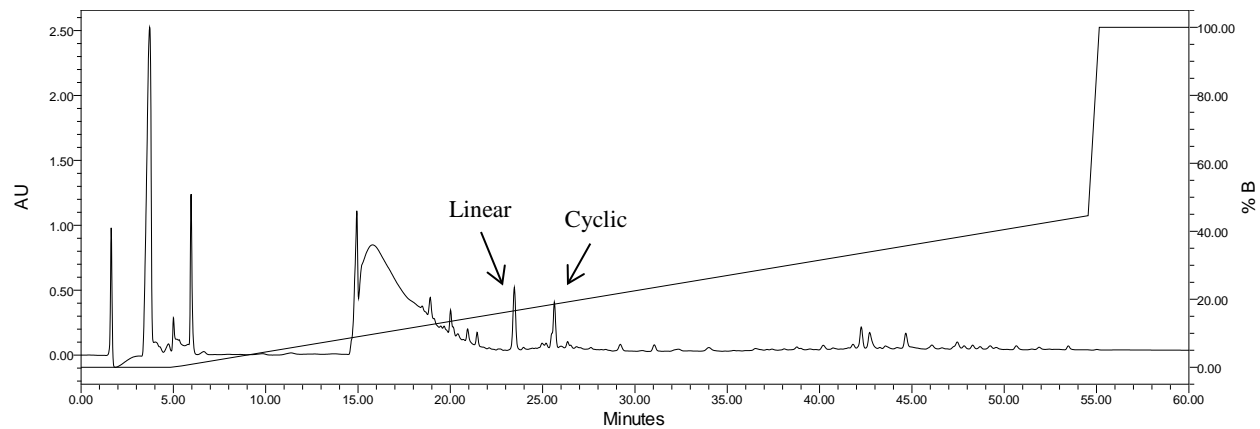
Peptide 9: cyclo(TKYRRE)BBNRM\*

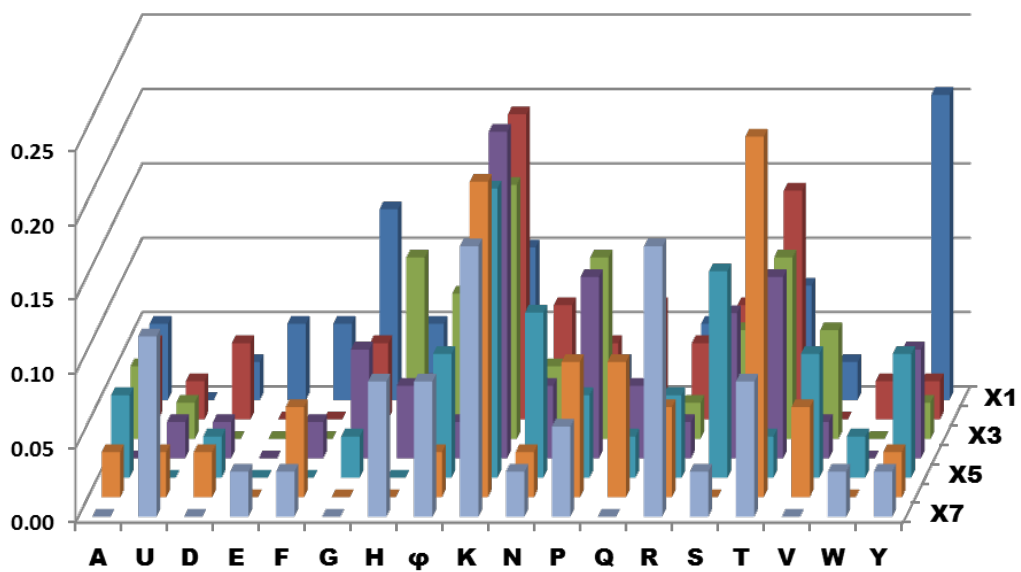


Peptide 11: cyclo(KVTYRRE)BBNRM\*



Peptide 13: cyclo(IKYKTNKE)BBNRM\*





**Figure S5.** Histograms showing the amino acid composition/sequence of poorly cyclizing peptides selected from libraries V when cyclization reaction was done in “Magic Mixture”. The y axis value represents the frequency (maximum value 1.0) at which each amino acid (x axis) was selected at a given random position ( $X^1$  to  $X^7$  on the z axis).  $\phi$  = Ile, Leu and Nle. U = Abu.