# Supporting Information

for

# Discovery of Unguisin J, a new cyclic peptide from *Aspergillus heteromorphous* CBS 117.55, and phylogeny-based bioinformatic analysis of UngA NRPS domains

Sharmila Neupane‡1,2, Marcelo Rodrigues de Amorim ‡1,2,3, and Elizabeth Skellam\*1,2

**Spectroscopic and spectrometric data of 1 and 2. Bioinformatic data of the biosynthetic gene clusters.**

Address: 1Department of Chemistry and 2BioDiscovery Institute, University of North Texas, 1155 Union Circle, Denton, TX 76203, U.S.A; 3 Instituto de Química de São Carlos, Universidade de São Paulo, CP 780, CEP 13560-970, São Carlos, SP, Brazil

Email: [Elizabeth.skellam@unt.edu](mailto:Elizabeth.skellam@unt.edu)

\* Corresponding author

‡ Equal contributors

**Table of Contents**

[**Table S1:** DNA and protein sequences of ung” BGC mined from *A. heteromorphus* CBS 117.55. 3](#_Toc152663781)S

[**Figure 1S:** A domain sequence alignment comparing UngA, UngA’, and UngA’’. 14](#_Toc153463471)S

[**Figure 2S:** C domain sequence alignment comparing UngA, UngA’, and UngA’’. 14](#_Toc153463472)S

[**Figure 3S:** A domain AMP binding motif identified in extracted A domains. 14](#_Toc153463473)S

[**Figure 4S:** C domain active site identified in extracted C domains*.* 15](#_Toc153463474)S

[**Figure 5S:** 1H NMR spectrum of **1** (500 MHz, DMSO-*d*6). 16](#_Toc153463475)S

[**Figure 6S:** Expansion of the 1H NMR spectrum of **1** (500 MHz, DMSO-*d*6). 17](#_Toc153463476)S

[**Figure 7S:** Expansion of the 1H NMR spectrum of **1** (500 MHz, DMSO-*d*6). 18](#_Toc153463477)S

[**Figure 8S:** Expansion of the 1H NMR spectrum of **1** (500 MHz, DMSO-*d*6). 19](#_Toc153463478)S

[**Figure 9S:** 13C NMR spectrum of **1** (125 MHz, DMSO-*d*6). 20](#_Toc153463479)S

[**Figure 10S:** *g*HSQC spectrum of **1** (1H:500 MHz; 13C:125 MHz, DMSO-*d*6). 21](#_Toc153463480)S

[**Figure 11S:** *g*HMBC spectrum of **1** (1H:500 MHz; 13C:125 MHz, DMSO-*d*6). 22](#_Toc153463481)S

[**Figure 12S:** *g*COSY spectrum of **1** (500 MHz, DMSO-*d*6). 23](#_Toc153463482)S

[**Figure 13S:** 2D-NOESY spectrum of **1** (500 MHz, DMSO-*d*6). 24](#_Toc153463483)S

[**Figure 14S:** HRESIMS spectrum of **1** in positive ionization mode. 25](#_Toc153463484)S

[**Figure 15S:** HRESIMS/MS spectra of **1** and fragmentation patterns in positive ionization mode. 26](#_Toc153463485)S

[**Figure 16S:** UV spectrum (photodiode array, H2O:MeCN) of **1**. 27](#_Toc153463486)S

[**Figure 17S:** HPLC-PDA chromatogram of unguisin J (**1**). 28](#_Toc153463487)S

[**Figure 18S:** 1H NMR spectrum of **2** (500 MHz, DMSO-*d*6). 29](#_Toc153463488)S

[**Figure 19S:** 13C NMR spectrum of **2** (125 MHz, DMSO-*d*6). 30](#_Toc153463489)

[**Figure 20S:** UV spectrum (photodiode array, H2O:MeCN) of **2**. 31](#_Toc153463490)

[**Figure 21S:** HRESIMS spectrum of **2** in positive ionization mode. 32](#_Toc153463491)

[**Figure 22S:** HPLC-PDA chromatogram of unguisin B (**2**). 33](#_Toc153463492)

[**Figure 23S:** HPLC-PDA chromatogram of soluble-organic extract from rice culture of *A. heteromorphus*. Peak labels represent the Unguisin J (**1**) and Unguisin B (**2**). 34](#_Toc153463493)

[**Figure 24S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of the hydrolysate of **1**. 35](#_Toc153463494)

[**Figure 25S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of the hydrolysate of **2**. 35](#_Toc153463495)

[**Figure 26S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of the D-alanine. 36](#_Toc153463496)

[**Figure 27S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of the L-alanine. 36](#_Toc153463497)

[**Figure 28S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of the D-valine. 37](#_Toc153463498)

[**Figure 29S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of the L-valine. 37](#_Toc153463499)

[**Figure 30S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of the D-phenylalanine. 38](#_Toc153463500)

[**Figure 31S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of the L-phenylalanine. 38](#_Toc153463501)

[**Figure 32S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of the D-leucine. 39](#_Toc153463502)

[**Figure 33S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of the L-leucine. 39](#_Toc153463503)

[**Figure 34S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of the D-tryptophan. 40](#_Toc153463504)

**Table S1:** DNA and protein sequences of ung” BGC mined from *A. heteromorphus* CBS 117.55.

|  |  |  |
| --- | --- | --- |
| **Gene** | **Putative DNA sequence** | **Predicted protein sequence** |
| ungA’’ | ATGCCGACAGAGAATGGGGAGGTCCAGCACCGGATCGACGGCCTTCAGGATCAGGATACTGTGGAGATATTGCTGGTTACCTTGGAATCATGGGCCCTGCAGGCTGTGGAGGAGTTCAAAGCCAGCAAGGGCATGCTCGAGATAGGCAGCGTGAGCGACAGATCGAACGATGCCCGTAAGGTGAAGATTGAGCTAGATGGGGGAACTGAGAATATATTGGCGCAATTGCGAGGGGAATGGGTGCGCCAGCTGGTGCACCAGCACCACGAAGCAGCCAGTAATTTTGCGGCGTCCATAATTTTGGATGGCAGCCCAGAGAGGCTGCTGGTCTTGCGGTTTATGGAAAGGAGGGAAGCCTGGGTGTCGGAGGAGGCGCCTAAGAGCTGGTGCCAGGAAAGGAGGTGTGCACCGTTCCTAGAGATGCATTGCCAGGCTCGCCCAGAAGAGCTGCAATTCCATATCGTCTCGCCTCCTCCTGTTGCGGCGCAAGCATTTGTGCAGCACCTGGATTTGCGGATCCGACAGTCCATTTCCAATCGCAGTCTCGCAGACGTGCCGGTAGAGACTGTGCCTGCCGCGTCGGACAGCGAGCTCAGGGCCATCTGGGGCTGGAATGGCACGCTTCCTGAGGGAGTGCAGAGATGCGTTCACGATATCATCGCCCAGCAGGCGCGACTGCGGCCACACGCGCCAGCAGTCTCTGCCTGGGACGGCGAGCTGGACTATCAGCAACTGGATCGGATCTCCACTCGATTGGCCCATTACCTGGTGCAGCTGGGTGCTGGCCCAGATGATATCATCCCCCTCTGCTTCGAGAAGTCCAAGTGGATGATGGTGGCCATTCTGGCCGTGATGAAGTCAGGGGCTGTCATCGCTGCCCTGGACCCGACACAGCCCGAGGATCGCCTGCGAAGCATTGTGAAGCAGCTGCAGCCACGCTGGATTCTGACATCGCCGGCCCAGATCGAAGTGGCCGCACGCCTGGAGATCAGCAATGTCATCGCTCTCGACGAAGGCCGCCTACGGCAACTGCCAGACTCCGAAGGAAGACATCTCCCATGCGTTGATCCCTCCCGTAACCTGTACATTGTTTTCACGTCGGGCAGCACTGGAACGCCCAAGGGAGTCATGATCAATCACACCAACTTTAGCAGCGCCATCGCCTACCAGCATGAGGCATTGGGGATGGATAACACCGCGCGAGTGTTCGACTTCGCTTCATACGCCTTTGATCTGGCCTGGGGGAACATCATCCATACCCTCGCGGCGGGGGGGTGTCTATGCATCCCCAGTGAAAGTGAGCGCCGTGGAAACATCGCCGAGGCAATCCGCAGACTTGGGGTGAACCATCTCCAGCTGACCCCCAGTGTTGCTCGCCTCATCGATCCGCGAGATATCCCCGCAGTGCGATGGATCCTCCTGATCGGAGAGCCAATGACCCAGGCCGATGTTGCTCAATGGACGCCATACTGCAAACTGATCAATTCCTACGGCCCAGCGGAATGCACGGTGGCGGTGACCTTCCAGACAATCCCCCACGGCCGACCTTGGGACTCGAGTATGGGGAAAGGCGTCGCGTGCAGTACCTGGATCGTTGACGAGGAGCATGGGGAGACACTGGTCCCCCTTGGCCACACAGGCGAGTTGTGGCTGGAAGGACCGCTGGTCGGCCAGGGATATCTAGGAGACCTGGAGAAGAGCGCGGCGAGTTTCATCGATAGCCCTGCCTGGCTGACCCGGGGAATACCAGGGGTCGTCCCCGGCCGGCGAGGCCGACTATATAAAACTAGAGATCTGGTGCGCTACAACCCGGACGGCTCGCTGGTTTACGTGGCTCGCAAGGATACGCAGATCAAGATCCGGGGCCAGCGGGTGGAACTGGGCGATGTGGAATATCACCTGAAGCTTGCACTGCCCGACAAGATTCCATCTGTTGCAGCCGAGGCCATCACCCCGCGGGGTAGCAGCAGCACCATCCTCGTCGCATACCTGGCCCTCGGCGAAGAAGCAACTGGCGCCGCGGAGAGCACCCGGGAGAGTCTAGCCAGTTGCCTGCATGGGGTGGAAGAGTACCTGGCCGACCGACTGCCTCGCTACATGGTGCCCAGCCTGTATCTTGCAGTCCCCGAGATCCCCATGACTGCCACCGGCAAGACAGACCGGCTACGCCTGCGTGAGATCGGGTCTTCTTTAACCCTGGATCAGCTGGCAGCCTTGCAGCCCTCGCGAGCAGTGGAGATCAAGGCTCCCGAGACCGAGATGGAGCACCGCCTTCAACAGCTCTGGGCGGCAACATTGAATATCAGCCCGAGCAGCATAGGGACAGGCGACAGTTTCCTCCGAATCGGCGGAGAATCGATGGCAGCCATCCGCCTGGTGCAGCTCGCCCGGAAAGAGGGCATAATCCTGACAGTGGCCGATATTTTCAACCACCCGCGCCTCGGCGATATGGCGCAAGAAGCCCAACGAGAACGGGCCCAGGACGTGACGACCATCCCTCCCTTCTCCCTGCTGCGTAGGGGCCACGACGTAAGAGATGCCTGCGCGTTGGCTGCAGCCGAATGCCGCGTTTCGGCAAACTCCATCGCCGACCTCCTGCCCTGCACCCCGCTGCAGGAAGGGCTGTTGGCCCTGACTGTCAAGCAGGCCGGGGAATATGTCCGCCAGATGGTATCAGAGCTTCCAGTGGATGTGGATCTGGCGCGCTTCCGCGCTGCCTGGGCCCAAGTGATTCATGAAGCGGCGATTTTGCGAACGCGGATTGTGGACCTGCCCAACGCCGGGCTCCTGCAGGTCGTAGTCGCAGACCAGCCCGAATGGGCCACGGGGAGCGATCTCGGTCATTTCTTGGAGTCAGAGAAAGGAAAGCCCATGGGACTGGGGACGTCGTTGGCGCGGTTTGGTCTGGTGTCGGATCAAGCGCAAGGCAAGGTGCTTTTCGTCTGGACCATCCACCACGCCTTGTATGATGGCTGGTCGCTGCCGGCGATGCTGGAACGGGTGGAAGCCATTTATGCGGGTGGCAGCTGCGATATGCTCCCCTCGTTTGCGGGGTTTGTGAAATATCTGGCGGATGGTACAGTGGAAGATGCGCAGGGTTACTGGCAATCCCAGTTCAATGGGATCCAGGCCGCAGTCTTCCCAGCCTTGCCCTCCCCAGACTACCAGCCCAAGTGTCAGGATCTGCTCCAATACCACGTCGCGAGTGTCTCCTGGCCCGGGAATGATATCACGGCCTCCACTGCAGTGCGAACAGCATGGGCAATTGTCGCAAGCCGTTATACCCAGTCTCCGGATGTGATCTTTGGGGCCACAGTGTCCGGGCGACAGGCACCGGTCCCGTTCGTTGAACGCATGGCCGGTCCAACCATTGCGACTGTTCCAGTCCGCGTGAATGTGCAGGGGGATGCCACGGTGGCCAGTTTGATGCAGAGCGTTCAGACACAGGCCGTGGCTATGATTCCGTACGAGCAGACGGGATTGAACCAGATCCGGCGGATCAATTCGGATGCAGACCAGGCAACCCAGTTCCAGTCCCTGCTGGTTGTACAGCCTCCAAGCTCGAAGACCTCTCGTCGGCCGGACGAGTGCTTGTTCCGAGTGGATCTGGATGCTGGTGACGAATTTCGCAGTATAAACACTTATTCGCTTATGTTGGAGTGCCATCTCGAGTCCGACGGGATGCGCCTGCGCATGAGCTATGATCGCCTGGTCATCGACGCGGAGCAGGTTAAGAGGATAGCCAGGCAGTTTGAAGGGGTGCTGCGCCAGGTCTGCAGTGAAGTGAATGCGCAGGAGCTGGTGGGCACTGTCTCAGCTGCCAGTGAAGCCGACCTGGCCCAGGTGTGGGCCTGGAATGCAACTGTCCCCCAGACCGTCCAGGGCTCTGTCAACGACCTCATTGCGCAACGCGTTCAGCAACAGCCCGATGCCCTGGCGATCTGCGCGTGGGATGGACAGCTGAGCTACCAGGAACTCGATACCCTCTCCACGCAGCTGGCCCTATCGCTCGTGCAACAGGGGGCAGGCCGTGGCTCGGTGATTCCCTTATGCTTCGAGAAGACAATGTGGACGCCCGTGGCCATGCTGGCTGGCATGAAGGCTGGCAGTACTGTGGTTACAATGGACCCCAGCCAGCCGGAGGATCGATTGCAGTCAATCATCCAACAGACCCAACCGCCACTCATCCTCACCTCCGACACATACAAGCCTCTGGCATCTCGACTCGCAGGGGTGGTCATCCGCGTGAATCGGGATACCCTGCAGGCACTGCCACAGCATGACCTTGATACTCCAGATCTCCCAACCATCCACCCGACGGACGGACTCTACATCGCGTTCACCTCGGGGAGCACTGGAAACCCCAAGGGGGCAGTCATGTCCCATCAGAACATTCGCAGCGCAATCCACCACAGTACGGGCGCCTTGGGCTTTACGCCTACCACCCGGGTGCTTGGATTTTCCTCCTATGCATTTGACGCGGTGTGGGTGGAGTTTCTCTATGCGATGGCCGTCGGGGGCTGTCTGTGTATCCCTTCGGACGCGCAGAGAAACAGTGGTGATCTTGCAGGCTGCATGGAAGGGTTGCGAGTGGACCTTGCTCTCCTGACCCCGTCCACGGCCCGGCTTCTCGACGCAGAGGCAGTCCCCAGCCTGAAAACCCTGGTCCTGATCGGGGAGCCCGTCACGAGCGATGACCTGGCCAGATGGACCGGCAACGTTGATTTGAGGAATGGGTACGGCCCTGCAGAATGCTCCGCAATAACCGCAGCCTATAAATTTCAGGGTCCAAATGACCAGCCCAGTATCGTTGGCCACGCGGTGGGCCTCGTGGCTTGGGTGGTTGAACCATCGGACGGCGCCAGCCTCAGCCCCCTCGGAGCGGTGGGGGAGCTGTGGGTGGAAGGGCCGTTGGTGGGAAAGGGGTATCATGGGGATCCAGAGAAGACAGCCATGAGTTTTGTCCATGACCCCGTCTGGCTGCTCCAGGGAACACTGGGCTATCCTGGGCGGCCAGGTCGACTCTACAAGACCGGCGATCTTGTTCGATATACCTCCGATGGCAAGCTGGTTTGCGTGGGGAGAAAGGACACGCAGGTCAAGATCCGCGGCCAGCGGGTCGAACTGGCCGAGATCGAGCATTATATCAAAGAGGCCACCAGAGCTTCCGTCGTGGTGGATATGGCCAGGCCCCAGGGCAGCAGAGGCCCCGTGCTGGTGGCCTATGTGGCTTTAGGGCAAGTTGCAGCCGTGGGCTCCCCCGAGGCAGTGAGGACGGCACTTCGGCGCTGCATCCAGGGGGTTGAAGATCATCTGAGCAAGCACCTGCCCCGGTATATGGTGCCGAGTTTCTATATCCCCGTGGTCGATATCCCATTAACAGCCACGGGCAAGACCGACCGACGGGCATTACGCAATACCGGATCCTCCTTCGCCATCGACCAGCTGGCAGAGCTGCAGCCCTCTGGAGGCCAGAAGCGCACCCTGCCCCAGAGTGAGATGGAGCGGAGCCTGCAGCAGCTCTGGGCGGAGGTGTTGAACATCAATCCCTCTCAGATCGGCATGGATGAGAGCTTTTTCCTGCTCGGGGGGGACTCTATCTCGGCCATGCAGGTCTCTGCCAAATCTCGCGCGAGGGGGGTTCGCGTCACAGTGCCAGATATCTTCAAATTCAAGACAATTGCGCGTCTAGCACGACGCGAGATCCAGATCGATGACCTGGCCATTGACGACCGAGAGATGCTCGACACCCCCTTTGCCCTATCTCCCATCCAACAGTTTTTCTTCGACGCCCAGAAAGACAAGCAGGGTCACTTCAATCAAAGCTTCCTCGTGCGCACTACCAAGCCCCAACAGCCCGACGCTGTACTGCGCGCCGTGCAATTCATTGTTGCCCGACACTCGATGCTGCGTGCTCGCTTCCACCGGGGCACCGATGGGGTGTGGACGCAACAGATCACCTCGCGTGCTGATGAATCCTATGCTTGCCGTCATCATCGCCTGGCTTCGGTGGAGGACGCCGTTCCGGTGCTGAATAGCAGCCAGCGGTCCTTGGACGTCCAGGCTGGCCCCATATTTGCGGTTGATCTGATCGACACTCCTCCAGACTGTCAGTACCTGTTTCTCACGGCGCACCATCTCGTTGTCGACCTTGTCTCCTGGCGGATTATCCTGGAGGAGCTGGAGGAGTACCTGTTGACTGGCACTGTTGCTGGTTCCTCTGCGCCACTCTCATTCCAGACCTGGTGCCGGCTTCAGGAGAGACACGCGCAGGACCATCTGACCCCCGGTACTGCCTTCCCGTTCGATATTCAGCCCCCGTGTGAGGCGTACTGGGGACTATCGCCGGAGCTCAATACGCACGATGCTATCCAAGAGGCCGGGTTTACCCTCAGCAAGAAAACGACCAGCATACTCCTGGGACCAGCGAACCATGCGTTCCAAACTCAGCCCATCGAACTGTTCCTAGCCGCGCTGCTGCACTCTTTCATGGACAGCTTCACTGACCGGGATCCTCCTACCATCTTCAACGAGGGGCACGGTAGAGAGCCATGGAACTCAGCAATCGACCTATCCCGGACGGTTGGCTGGTTTACCACCCTGGCGCCAATCTCAGTGGAAGCGGGCTCCCATCACCTTAGCATCCCCCAGTTCGTCCGTTACATCAAGGATCGCCGGCGACAGATTCCTCACAATGGGTGGTCGTACTTCACCTCACGGTACTTGAACACCGAGGGGAAGACGGTCTTTAGCCGCCATGCCTGCCCTGAGGTTACGTTCAACTACTTTGGCCTCTATCAGCAACTGGAGAGGAGCGACGCGCTCTTCCAATCGTGTCCCAGTCTGCAGGACCGAGTGCTGGATGTTGCAGACCGCATGGCCCGGTTTGCCCTGATTGATGTTTCTGCCGAGGTCACGCAAGACTGCCTGCAGTTCCGGTTTCTTTCCAGCCAGCGCATGCAGAAGCAGGACGCCCTGGCTCGGTGGATTGCGGCGTGCGAGCGTTTGTTAGAAGCCACAGCCGTGCAGCTGCTAGATATGCAACCCAGCTACACGTTGGCCGACTTCCCATTGCTGCCCCCTACAGATGCAACGTGGAAGCAGCTGGAGACGCTCCCACGGCTTGGCCTTGCCTACGGTGACGTGGAGGATATCTATCCATGCTCGCCCTTGCAGCAGGGGATTCTGTTGAGTCAGGCAAAAAGCCCGGAGATGTATTGGACCCGAGTGCGTTGGGTTGTCCGGTCCAGTACGGGGGCATCATCTCCGGTTGATGCAGACCGACTGGAACGTGCCTGGCGGAAGGTCGTCGCACGCCATGCTGCACTCAGGACTAGATTCTCTGACAGCCCCTTCTCAGATGGATATAGTTACCAGATAGTCTTGAAAGACCCTCCCCCCACTATTCACACCATCCAGGCTGCGGACCCGATGGACGCCGTTGTCGAGTATCGAGCAGCCAGCGGGTCGCAGGTCCGGCCAGTGCACTCTCTGGTCCTTTGCCCTATGGCCAGCGGAGACCTCGTGTGTGACCTGCAGATCAACCACGCCATCATCGATGCCATATCCATCCGGGTTCTGAAACGCGAGCTGTGTGCCGCGTACGACGACGCCCTCCCTGCGGACCCAGGGCCCCGGTACAGTGACTATATCAGCCATCTCCAGTCCTTGCCGACGACGGAGGCGAAAGAATACTGGCAGACGCAGTTGGCTGGCGCTCAGCCATGCATCTTCCCTACCCTGAATGAGCCCATCGCCGAACCCAAGGACGCGGCCGCCGTTACATCTCTGCCGATCAGCCCGGAGACGGACCAGGCCTTGCGCCAGTTCTGCCGAAGCCATGCCCTGACCCCGGCCAACGTGTTCAGTCTCGCGTGGATGTTGGTCCTACGCAGCTACACCAGCAGCGAGTCCGTCTGCTACGGCTATCTGATCTCTGGCCGGGATGTGCCCATACACCAGGTAGATCGAGCGGTTGGACCGTTCATTAACATGGTGGTCAACCATGTCGAGATAGACAGCAACCGAACCCTCCTAGCAATGCTGCAAGAGGTTCAGGCAGGGTACCTAGGAAGTCTGAAGTACAACCAATATCCGCTGGCCGAGATCCTGCATGACCTCAGCACCGAGGGGCAGCCCTTCTTCAATACGGTTCTGTCGGTGCAGAGTGGCGGCAGGGCCGACTCAGCCCAGCCCCCAGACACACCCACCATCACCCTCGAGAACGAGACGTGGCACGATCCGAACGAATACGACATTGCCGCATCGGTCCTGCTGCTGGACGACAATCCCAGAATCACGCTGAACTACTCCCGCCATCTGCTTTCGGAGCGACAGGCCAGTGCCGTCGGGGCGACGTTTGTGGAAGCCCTGGCCAATATCGTCCGCCATCCGGACAGAAACCTCTGCGATCTGGAGACGACGATCAGTCCCCAGGATCTCGCCACCATCTGGGACTGGAATGCCAAGGTTCCCGCGACCTTGTCCAGCTCGGTGCCCGAGCTGTTCTCCCGATGGGTGAAGCAACAGCCTGATGCGCAGGCCGTCTGTGCCTGGGATGGAGACTTGACCTATCGTCAACTCGACGAGGGGTCGTCCCGGCTGGCTCACCATCTGCTGGCTCGGGGTGTCCGCCCGCAGTCAATACTGCCCCTGTGCTTCGAGAAATCCCGGTGGGTGCCCGTGGCCATGCTGGGGGTGCTCAAGGCAGGCTGCGCGGTAGTCACGATGGATCCGGAACAGCCCGTGGAGCGCCTACAGCTCGTGGTGCAGAAGACTCAAGGGGTGATTCTGACATCGCCAGCCTGTCAGGGCCTAGCATCCCAGCTACGTCCTGAGGCTGTCGTCGTGGATGGGCGATCGCTTGAGAGGATGCCTCAACCAGACCCGCTGTCTCTGCCCACGATCCATCCCACGGACAGGCTCTTTCTGGTGTTTTCCTCCGGCACCACAGGGACCCCCAAAGGGTCTGTGATGAGCCATCAGAACGCATGCTCGGCAGTTCACCACCAGCAGGCACGTCTTGGGCTGCCGCCCGCAGCCCGAATGATCGATTGCCTCTCATACGCGTTCGACGCTCCCTGGTTCAACTTTTTGCATGCCTTTAGCTCAGGCGGTTGTCTGTGTGTTCCGTCCGATCAACAGCGGAAGGACAATCTGGCGGGCTGTATCGAGAGCCTCAAGGCCAACTACGCGCTTCTTACGCCCTCGATTGCCAGAGTCCTGGACCCTGCAGCGGTGCCTGGCCTGGAGACACTTGCATTCGGCGGGGAGGCTATACGCGCGGAGGATATCAGTCGATGGCACGAGGTCAAGCTGATGGGGTATTATGGTCCCTCGGAGTGCACTGTCGTCTCAACCATCCACCAGTTCGAGAGCAAAACAGAGGAACCCAGAATGCTCGGTTATTCTTCTGGGCTGACAGCCTGGGTGGTGGACCCGTTGCAGGGGAGACGCCTGGCACCGCTTGGTGCGACAGGAGAGCTCTGTGTCGAGGGTCCGCTTGTCGGACAAGGGTATATCGATGAGCCCCAGAAGACTGCGGCCAGCTTTGTGGAAGATCCCAGCTGGCTTCTCCGTGGTGGAGGGCCAGACTTCCCTGGACGTCGGGGCCGGGTGTATAAAACAGGAGATCTCGTCCGATTTCGTCCCGATGGATCGCTCCTGTATGTCGGGCGCAAGGATACTCAAGTGAAGATCCGGGGCCAGCGGGTGGAGTTGGGCGAGGTTGAGCATCACATCCGCCAGACGCTGCCAGCCGGCGTGGATGTGCCGGTGGTTGTGGAACTGGTCCACCCCCATGCAAGTGCGAATCCGGTGCTGGTGGCCTTTTTGCCCGTGGGAAAGGAAGCTGCTGGGTCATCCGAGTCAATGCGTGCAGTCCTGGACCGCTATACGCACGGGGTGGAGGACCGCCTGATGGCCCGACTGCCCAGCTATATGGTCCCTCGCATGTACATCCCCGTGCCAGCGATCCCATTCACGACGACGGGGAAGACAGACCGGCGCAGCCTGCAACGGGTCGCGTCTTCGTATACCCTGGAGCAGCTGGCAGCGCTTCAGCCCTCGCGGGACACTCGTCGTGCGCCCACAACGGCGATGGAGCGGCGTCTGCAAGGGATGTGGGCTACTATCCTGGATATCGATGCAGTGACCATCGCGGCCACCGACAGCTTCCTGCGGATTGGAGGTGACTCAATCGGAGCCATTCGACTCGTGCGACTGGCTGCAGAGCAAGACATTCTGCTCACGGTGGCAGCCATCTTCAAATCTCCCATTCTGTGTGACATGGCGCAGGTGGCGACACTGGGATCGACTTCCGGACCTCATGACATCCCCCCATTTTCACTGCTCAAGCAAGACGTGGATATCTCGCAGGCGCGATTGCAGGTGGCGGCGTGGTGTGAGCTCAGCCCGTCTGCTGTCGAGGATATCCTTCCGTGTACACCTCTCCAGGAAGGTCTGCTTTCTCTGACGGTCAAGAATCAGGGCGCCTATGTGAACCGCCAGGTCCTGCAACTTCGAGACGAGGTAGACCTTGGACGCTTCCAAACCGCCTGGAATCAAGTCGCCATGGCCACGTCGATCCTGCGAACCCGCATCGTGGACCTTCCGGGGCAAGGGTTGGTGCAAGTCGTTACCAGCGAGGTGCCACGGTGGAACCATGGCCGTAGCCTCGACGCACTCGTGCAAGACGATCTGCAACGTCCGGTAACCCTTGGTACCCCCTTGGCCACATTTGGCTTGGTGGACGCCGGTGACGTTGGTGGGGATCGGCAACGGTACTTCCTGCTAACGCTTCATCATGCGCTCTATGACGGATGGTCTCTGTCGCTGCTGCTGGACGAGGTCAGCAAGGCCTACCACGGGACCGCGAGCGACGGCCTTGTCTCCTTCAAGACAATGATCAAGTATGTGACAGAGCTGGGTGCCGAGGCGGACTCCTACTGGCAAGGCAGCCTGGATGGTCTGGTGGCCGAGCCATTCCCTGCCCTGCCCTCTCCCATGTACCAGCCACGCGCGCAAGACATCCTGGAGCACAACGTATTAGGGCTACGGTGGTTGCAAAACCACCACATCACACCGGCAACCACTCTTCGCGCGGCCTGGGCCATTCTCACAACCCACTACACCCAATCTGCGGATGTCCTTTTTGGATCGACCGTGACTGGACGGCAGGCTCCCATTCACCGCGTGGGACTGGTCGAAGGGCCGACAATTGCAACTGTGCCGGTCCGCATTGCCATCCAGGAAAAGGCTACCCTGGCAGGCCTACTCGACCAGGTCCAAGAGCAGTCTGTCGATATGATTCCCTACGAACAGGTCGGGTTGCAGCGAATCCGGAGGCTCAGCGCCGACACCGAGCGAGCTTGTCAGTTCCAGACGCTCCTGGTGGTGCAGCCAGCCCCGGAACCGAGCGCCACAGATGGACCCTCTGCACTGTTCCATGCAGAAGATGACACTCTCAGCCAGGCGGCTCTCAGCAGCTTCAATTCATACGCTCTCTTGCTGCAATGCCAGCTGACGGCCACGGGGGTGTCCATGCAGATGAGCTACGACTCGCATGTCATGGCACAACCGCAGGTGCAACGGCTGGCCCGGCAGTTCGAGCATCTAGTACGACTCCTTTGCGATGAGTCTCAGCACTCATCCTCTGTTGCTCACATAGATGCGGTCTGCGAGGATGATCTGCGAAGGATCTGGGCGCTGAACTCCCCGCCCCCGGCCAGAGTGGACGCCTGCATGCACGACTTGATCACCGAGCAGGCCCAGCGGCGGCCAGTCTCACAGGCCGTCGCTGCATGGGATGGAGAGCTCACGTATCGCGAGCTGGACGAGCTCTCCACCCAGTTGGCATATGGCCTCGTCAACTTGGGGGTCGGCCCCCATGCGGTGGTCGCTCTGTGCTTTGAGAAATCAATGTGGATGCCCGTGGCAATGCTTGGAGTCATGAAAGCCGGCGGTGCATCGGTGGCCATGGATATCACGCAGCCAGAGGACCGGCTTCGCACGGTAGTGCAGCAGGTACAGCCTCCTCTGGTCCTGAGCTCCCTGGAGGCAGAGGACCTGGCACGGCGTCTGAGCGATAGGCCGGTGCACGTTGTCTCTCAAGACAGACTGCAGACGACGTCGCTGGGCAGCGGCCAGGGCGGCCAATTGCCCCCAGTTCAGCCGACAGACAGGCTGTATATTGCCTTCACCTCAGGTAGTACGGGGGTCCCGAAAGGCGCGGTCATGTCCCACCAGAACTTCACCAGTGCCGTGCACCACCAAACCGCGCTCGAGTTTACTGCGTCGGCCAGGGTGTTTGACTTTTCGTCCTACGCGTTCGATGCTTGCTGGCTGAACTTTCTCCATACGATGGCGGCGGGTGCCTGTCTCTGCATCCCCTCCGAAGAGGAGCGCAAGAGCGACATCACTGGCTGCATGCGCCGCCTGGCCGTGACTTATGCGAACCTCACGCCATCCACCGCACGTCTGATCGATCCGACGTCCGTCCCCAGCCTCCAGACCCTCGTTCTGGTCGGAGAGCCGGTGGCACAGCAGGACATCACTCAGTGGAAAGCCCACGTCCAGCTGAAGAACGGTTATGGCCCTGCGGAATGCAGCGCCATCTCCACCACGTTCGATCTCGGGCAAAGTGACCACGACCCGGCAACCATCGGCGCCGGCAGAGGAATGGTGACCTGGGTGGTTGAGCCGACCGAAAGCCGACATTTGTCGCCGTACGGGGCTGTGGGCGAGCTGTGGGTGGAAGGTCCTCTGGTTGGATTAGGGTATCTCGGCCGTCCGGATCTCAGTGCTGCAAGCTTCGTCGACAACCCCCCATGGCTGCTGCGCGGGGGACCCCAGGGGTTCCCAGGTCGACATGGCCGGCTTTACCGCACCGGTGACCTGGTGCGGTATAATCTGGACGGCACCCTGGTCTGTATTGGCCGCAAAGATGCCCAGGTCAAGATCCGCGGACAGCGCGTGGAGCTGGCCGAGGTCGAGCATCATCTGCACCAGGCCTTACCTTCGGCCGCCGTGGACGTCTCTGTGGCAGTCGAGGTCATCACCCTGCAGGGCAGTGCCAATCCACTGCTAGTCGCCTTTGTGGCGATGGGGGAGGCGGCCCTGGGCCCGCGGGAGACTGTCCGAGCAAAGCTGGCTCTTTACAGCCAGGGGGCCAGGGAGCGCATGGCAGACCAGCTGCCCGTCTACATGGTGCCCAGTCTGCTTCTTCCGGCGGTTGAAATTCCCACGACGGCCACGGGCAAAAGAGATCGACGGCGGCTGCGTGAGATGTGGGCCTCCAAGTCCCTGGAGGAACTGGCCGAGCTGCAGCCGACCAAGGGCAACCACCAAGCACCGACGACAGACGTCGAACGGCGGATTCTGCAACTGTGGGCCGAGTGCTTGAATATCAGCGCATCGAAGATCAGCATTCACGATAGCTTCTTCGCCCTTGGCGGGGACTCAATCTCCGCCATGCAGCTCTCGGCAAAGGGCCGCTCCATCGACCTCCAGATGACTGTAGGGGACATTTTCAAGTACAAGACAATCGCTCGACTGGCTCTCAGCATCTCCCCAGCAGTGGAATCGGCCGTGTTACATGCCCCAGAGGGCCAGGGTGCCTCCTTTGCCCTGTCTCCTATCCAGCAGATGTTCGCTGACATGCAGCAGGGGGTTAGCAACCATTTCAACCAAAGCTTTTTCGTGCAGGTACGTCGACCGGTCACCCTGCCTCAGATCCAGGCCGCTGTGGATGCCCTGGTCGCCCATCATGGCATGTTGCGCGCGCGTTTCAAGTGCAGCAGGGAGAATATCTGGAGCCAGTGGATTCTGCCCCCGGGGACCACAGGAAGCTATCGCGTTTGCCAGCATGAGGTTGCGGGTCTTCAGGCAGCCTCGGCGGTGATCAATCACAGCCAGGAGTCGCTAGACATCCAGAACGGTCGGCTCATGGCGGTCGATCTGATCAATGCAGACGAGGGCCAATACCTTTTCTTGGTGGTCCACCACATGGTCGTCGACCTGGTGTCTTGGAGAATCATCCTGGCTGATCTGGAAGAGCACTTGACCACCGCATCCCTCTCCGGGTTCACCTCGATGTCCTTCCAGACCTGGTGCCAGTTGCAGGCTGACCACCACGCCGAGAGCCCCTTCGAGCTGGAGGCTGTCCTGCCAAACGGGGCTCCGCCTCCGCTGCTGCCACAGCTAGACTACTGGGGTCAGGTCCGCAGCTTAAACACATTCGATAACATGATCAAGGATGGCGTGGTTCTCGGCAGGCAGGACACAGAGACCTTGCTAGGGCCTGCGAATGCGGCATTTGACACACAAATGGTGGAGCTCCTCCATGCGGCCATCCTACATTCCTTTGCCAATGTCTTCCACGACCGGGCACCTCCGACGGTATTCAGCGAAGGCCATGGTCGGGAGCCCTGGAGGTCTGCAATCGATATTTCACGCACGGTCGGGTGGTTTACCACGATGTTCCCTGTGGTCGCCACAGCCAAGAAGGGAGATAGTATTGCCAGCATTGTGCGCCATGTCAAGGACCGTCGACGCCAGATACCCGATAATGGCCGGCCATACTTTGCAAGCCGCTTCCTCACTCCGGCTGGCAAGCGGGCGTTCCAGGTCAATGGCCCTGTCGAGGTGATTTTCAACTACCTCGGGCTGTACCAGCAACTCGAGCGGTCGGACGCCCTCTTCCACATGCGCGAGATGCCGGACGGGGTGGATGACATGGCTGATATCTCGGGGAGCCTGTTCCGATTTGCTCTGGTGGATATCTCGGCATCGGTGACGGATGGCTGTTTGCATGTTGACTTCATGTACAATCGGCACATGCAACACCAGGCCTCCATCCGCGCGTGGATCAAGGAATGCAGGCAGTCCCTCCACGCCGCAGCGCAGGAGCTTCCTCTGATACAGCCAAGTTATACCCTCTGCAGCTTCCCGCTGCTGCGCATGACAGACCCGGCATTGGCAATCCTGCAGCAGAGGCTGACGGAGCTGGACCTCGCATACGGGCAGGTCGAAGATATCTACCCCTGCTCGCCGCTCCAGGATGGAATCCTCATGAGCCAGATCAAGAACCCAGACCTCTACAGGACACGCATCCGGTGGATGGCACAGTCTGCACAGGGATCAACAGCAGTTGATACCAACCGGCTCAAGCACGCGTGGCAGCAGGTGGTAGACCGCCATCCAGCCCTGCGCACCATCTTCGTCGACAGCATCTCGGGCAGGGGACTGAAGGATCAAGTGGTGGTGAGGAATCTCCGGGCAGACGTCCACATCGTCCAGTCCGCGGACGAAGCAAGTATCGAGGACATGTCCCCAACAAAAGGCAAAACCAGCTCAGTGCTCACTCTGTCGACGACGAACTCGGGCGTCCTGTGCGAGCTATCGATCAGCCATGCGCTCATCGACGCCTTTACCTTGGGAATCCTCAAGCAGGAGCTGTGTGCCGCATATACTGGCCTCCTTTCCTCTTCCCTGGCACCCCTGTATAGCGACTACATCCACTTCATCCAGTCCCTGCCCGCTGGGTCAGCCGCGGCATACTGGCAGGAGCATCTCCAGGAGGTCAAGCCCTGTCTATTTCCATCCTTGGGCGGTCTAAATACGGAAGGCCGGCGGTCGCAAGCTCACATCCCTATTACCTTCGAGCGAGAGCTACACCTGGCATTGCGCATCTTTTGCATGGAGCATGGACTGACGGTGTCTAATATCTTCCATGTCGCTTGGGCCCTGGTCCTGCGGGCTTATACTGGTTTGGATACTGTCTGCTTTGGCTATCTCACATCAGGCCGAGATATTCCTCTGCAAGGGGCTGACGGAACCGTTGGTCCATTTATCAATATGCTTGCCAGCCGCGTGGACCTGGGCAGTAAGGACTCGCTGATGACCCTGGTACAGAGGGACCAGGAGCAGTATCTCAACAGTTTGGAGTTTCAGCATTACCCGCTGGCCAAGATCTTCCACTTGATTGATACGCCCGAGAAGGGGCTTTTCAACACCGCCATGTCAGTCCAAGCTAGTGACTCTGGTCCCAAGGACTGCCAATCGGCTATATCGCTTGTAGATGAAGGAGGAGATGATCCCACCGAGGTAAGAAAGATAAGAAGACTTTTTCCTTTGTGTGACTGCGCTAATGCGAGACTCAGTATGATATCATGATAAACATTGGCGTCGGAGACGAGGACACCGGATGCAACTTTACCTTTAACGAATCGGTGATATCTGACCGCTATGCAAAGAGTGCCATCGACCTGTTCTTGCACGCCGTCTCACATATTGTTCAGCACGCGGACCAGACGGCCCAGGAAGCCAATTTCATCAGCAAGCAAGACCTGCAATCGATCTGGCAGTGGAATGCCGCCGTAGCAACGCCTCTGGACTGGTGTGTTCATGAACTTATCATAGAGCAGGCAGAGAAACAACCAGCGGCGGCTGCCATCTGCGCGTGGGATGGTAATCTCACGTATAAGCAACTCAACGATCTGTCCACGCAGCTGGCATGCCACATCCGACAACTGGGTGTTGGGCCAGGGGTTAATGTCCCGTTGCTCTTTGAGAAGTCGCGGTGGATGTCGGTTGCCAGTCTGGCGGTCATGAAGGCTGGTGGAACCATGGTGGGCTTGGATCCTGGTCAGCCTGCAGGGAGACTACGAAGTATCATCGATCAAGTCCAGCCTTGTCTCATTTTGACTTCGGCCGGAAATAGCAACAGCGCGGCAGCTTTTGCATCCTGCCATGTTGTCCGGGTGGATGACGCCAGTTTGGCCCAGCTGGCCACGCCGTGCAATGCTCTGTTACCCCCGGTTGACCCTGCCAGCAGCCTATACCTGGTATTCACCTCTGGGAGTACAGGCGTGCCAAAGGGTGTGGCCATCAGCCATTCGAACCTCAGCACCGCGATTACTCACCAAAAGCGCATCTTGAAACTTTCTGTCGCGTCGAGGGTACTTGAATTCGCTTCATATGCATTCGATGTCTCCTGGGGTACAATCCTTCACACCCTTGCGGCTGGTGGCTGCGTCTGCGTGCCTGAAGAGTCGGAACGCCGCGGAGATATCAGTGCCGCAATGCGTCGGATGGAGGTAAACTACGCGCATCTTACTCCATCTGTTGCGCGTCTTCTGAACCCTTCCAATGTGCCGCTCCTCCAAACCCTCGTCTTGAGTGGAGAAGCCGTATCGCGAGCGGATATGGAGCAGTGGAGCCGGCACGTCCATCTGATCAACGCGTACGGCCCTGCTGAGGCCGCCGTATGGGTGATATTCGCACACCTCGACTCAGCATTGTCCATGCCCTCCATCGGAAAGGGGGGTGGCTGCACCACGTGGATTGTTGACCCTAGTCGACCCGACCAATTAGCCCCAGTGGGCTGCACGGGTGAGCTCTGGCTGGAAGGTCCACTAGTGGGACGTGGCTATCTTCACGACCCCGAACGAACAGCAGCCGCGTTCATCGAGAACCCGCGATGGTTGCTCCAAGGAGCAGGAGGGGCCCATGGCCCTGGCCGACAGGGACGGCTTTATCGGACTGGAGACTTGGCCCGGTATACACCCGACGGTTCGATTGTTTATATTGGGAGAAAGGATAACCAGATCAAGATCCACGGTCAAAGGGTGGAGTTGGAAGAGGTAGAGAAATATATTGAGCAAGCCATGCTAAACAGCGCCGCAGCACCAGCAGTCCCTGTGGTAGCAGCGGTCGTCACCCCTCAAGGAAGCAAGAAGGCAATTCTGGCAGCTTATCTGGCTCTGGGTGAGCCGGCCACGGGCTCTGTGGAAATTGTGCGCAAATCTCTCAGCAGGTACACTGGGATCATCAACCCGGCCTTGGAAGAGAGCCTGCCGACCTATATGAGGCCGAGCATCTACATCCCCGTGGCGGAGATCCCCACGACTACCAATGGAAAGGCTGACCGCAGCAAACTCGCAACCATGGCATCCAGTCGCACTCTGGCTGAATGGGCTGGCCTTCAGACCTCTGAGGGCAAATGGCGCCCTGTATCGTCTCCGGGGGAGCTGGGGCTGCAGAGACTGTTTGCTGAAGTGCTGAATATGGACCAGAGCCTAGTAGGCATGGACGACAGCTTCTTCTCGCTCGGCGGTGATTCCATTACCGCGATGCAGCTGTCGGCCAAGTCCCAATCCAGCCTTCTCTATATCACGGTGGGCGATATATTCAAGCACAAGACCGTCGCACAACTGGCAAGCAATGCAAGGCAGACAGCGAGCACCTCCGTGCAGCTCCCGGAAGTCCCCAACAGCTTGTTCGAGCTGTCCCCGATCCAGCAGCTATTCTTTGCCTCGCAGGACAAGGGCAAGAACCTGTTTAACCAGTCCTTCCTGGTCCGCGTTTCCCGCTCCCTGAATCCCAACGAGCTCCAGAAGGCGATTGGAGTTTTGGCAGCACGCCATTCCATGCTCCGGGCTCGGTTCGTGCAAAGCGCGGACGGCGTCTGGCAGCAGAAGATTGTAAACGATTCAGCTGGCTGCTATACATTCCGCAGCCACCACATCACCACGCTGCAAGATATGGAGCCCCTCCTCCATAGCAGCCAGGAGTTGCTGGATATTGTCCAGGGTCCCATCCTGGCCGTGGAACTGATTGATTCGAGCCACGATGGGCAGTATCTCTTCATGACAGCCCATCATCTCGTGGTGGACCTTGTCTCATGGAGAATTCTGCTGGGCGATCTGGAGGAGTTCATCCGCTCGGGCACAATCACCGGATTCCCTCCATTCTCATTCCAATCATGGAGTCAACTGCAGGCGAACTATGCGCGAGACCATTTGCCTCCAAAGATCGCGCTGCCTTTTAAGGTCAGTCCGCCCCGGCATGAGTACTGGGGACTGGTCCCTGGGCACGATGCAAACACCCTGAGCGATTCCGGTCGAGGCAGCTTCACGATGAACAAGCGCCTGACCGACATTCTCATGGGCACGGCCAACTCGGCCTTTGATACGCAGCCGGTTGAGATCCTCCACGCAGCCCTGCTGTATGCGTTTGCACAGACATTCCAGGACCGCGAAGCACCGCCTCTCTTCACCGAGGGGCATGGTCGAGAAGCCTGGGACTCGGGCATTGACCTGTCCAGAACTGTTGGCTGGTTTACCACCATCTTCCCTGTGGCAGCTTCGATTACGCAAAACCACAGTCTTCCAGAGGTGGTGCGCCGTGTGAAGGATACACGCCGGCAGACACCACGCAACGGGTGGTCCTATTTCACCTCGCGGTACCTGAATACTGATGGACGGCAGGCCTTCCAGATCAAGGGCCCTGTGGAGATTATTTTTAATTACATGGGCTTATATCAGCAGCTGGAGCGGCCCGATTCCCTGTTCCAGCAGTGTGATATTGCGGTCACAGCGCCCCCTGCTGCTGCGGCGACTCTCTCGCGGTTTGCCCTGATCGACGTCGCAGCATCTGTTGTGCAGGGTCAGCTGAAGTTCGAGTTCCTATACAACAGGGAGATGAAAGGACAAGACAGGATACTTGAATGGATCAGCAAGACAGAATCCTCACTCAAAGCGGCAGCGGAAGAACTTCTGTTGCAGAGCCCAAGCTATACCATCTGCGACTTTCCCCTGCTCTCTCTCACCGACCAGGGTCTCGACGAGCTTCTCAATAGGGTCCTCCCGGCTGTGGGTCTTTCCTACGGCCAGATCGAAGACATCTACCCCTGTGCGCCGATTCAGGAGGGCATCATCATGAGCCAGGCCAAGAGCCCTGAGCTCTATTGGACTCGAGTGCGGTGGACGGTGCAGTCAACAGGCACGTCGCCAGTCGATCTCGATCGGTTGAGGCGGGCCTGGCAGATGGTCGTGAACCGTCACTCGATTTTACGCACGATATTTATTGACGGTATCGGATCCGGCCAGGTGAAGGACCAGGTGGTCCTCAAGGACTTTCCGGTGGACGTCGAGGTCCTCCATGCCGAGCTGCATCAGCTCGCAGAGCCAGTCACTGGGCGTCAATGGCACACCGATCTGTCGAGCAGAAGAGACACCCCGCAGCATTCCCTGGTGCTGACGCAAACAGCGTCTGGAGTCGTGTTCTGCGACTTGGAGCTCAACCATGCGATGGTCGATGCGTATTCTCTAGCCCTTTTGAGGCAGGAGATCTGTGCGGCATATACTGGCTCACTGCCTGCCACACCCGCGCCGGCCTACCAAGCGTATATTGAGCACCTGCAGGGGCTGTCCCTTGTTGAAGGACAACGCTTCTGGCAGACATATCTTGACAACGCTCAGCCCTGTCACTTTCCGGCCCTCGGCCAGCCGGATGTCACAGACGGCTCGAACGCCCGCCGGGCGCTATCCATCTCGTTGGACGCGGCCACTCATCAAGCCCTGCGTGTCTTCTGCCAGCAGCACGCAGTTACACCCTCGAACGTGTTCTACCTGGCTTGGGGGCTTCTTCTTCGCGCCTACACTGGCTTAGACACAGTGTGTTTCGGCTATCTTACGTCCGGGCGCGATGTCCCTGTCCCAGGGGTCGATAGGATTGTTGGTCCGCTCATTAACATGCTTGTCTGTGTTCTGGAATTCAGGACTGGGGCATCGGTGCGATCGGCCATGCAGAAGGTCCAGGAAGACTATCTCGCCGCCCTGCAGTATCAATCTACGCCTCTCAGCAAGATCCTGCAGCTTTCTGGGACGTCCGGTCGGGGCCTATTCAATACCGGCATATCGGTCCAGGGCGGTGCTACTTCCGGCGAACCGGACGAGCATGATATCATAGTGACGGATCAGACAGGGCTGGATTCGCCAGAGGTATGTGTTTTGTTCTTCTTGTCGTGGTCACGATTCCCCTTTGCTCTCCATAGCTGACCCGGTCCCACAGTACGACATTGCCGTTGCTATCTCCCACGATGAGGAGGAGACGGAGATTGCATTCGACTACATGGCCGTCACCTTGTCCGACCAAGGGGGTCAAAGCCTTGCAGGTCTTTTTGTTCAGGTTGTCGCTGATGTCATACGTGCTCCGGATCAAACGGTCCAGGCTGTCAATGTGATCAGCAAGGAAGACCTCCAGAGTCTGTGGACCTGGAATCGAACAGTCCCGGAAACCGTACAGGCGTGTGTTCATGATATGATTGGACAAAAAGCCTGTGATAGCCCTGATTCCCCGGCAATCCATGCATGGGATGGCGCGCTTACATATCAAGAGTTGGACCTTCTGTCCACCCGACTGGCCCGGTATCTTATAGATCTAGGGGTCCGGCAGAATACAGCCATCCCTCTATGTTTCGAGAAGTCCATGTGGATGCCGGTAGCTGCACTGGCAGTGATGAAGACAGGCGCTGCATGCGTGGCCATGGATATGACCCAGCCAGAGAAGCGCCTGCAGGCGATTGTGGACCTGGTGCAGCCGGATCTTCTCGTGACATCCGTCGCAAATCGCAAGACCGTCCAGCAACTGGCAGACACGAAGGCCGTGTTGGCGATTGACCAGGCATTCTTTTCCCAAATCGCACCTCCCACGTCGTCCTGCCTGCCGACAGTCTCCCCATCAAGTCCTCTCTATACCGTCTTTACATCTGGCAGTACCGGCACACCCAAGGGAGCAGTCATCAGCCATGCCAACTTTGCCAGTGCCATTGTGCATCAAACGGACCTGCTCGCACTGGACCCCGATTCGCGCGTCTTTGACTTTGTCTCGTATGCCTTTGATGTTTCCTGGTCGAACCTACTCCACACGCTGGCCGCTGGGGCATGTTTGTGCATCCCCTCCGAAGCCATGCGCCGGGACAACCCTGTCGAGGCCATGCAGGCGATGAAAGTGACCCATGCCCAATTTACCCCGTCGATGGCCCGTACAGTGGACCCGGATCAATGCAAGACGCTCAAGGCCTTGATCCTGGGCGGCGAGGCCATGTCCCAGCACGACATCGCTGTCTGGGCACCTTGGGTTGACCTTCGTGTTGCGTATGGTCCCGCCGAGTGTACGGTTGCCGCCGTCATGGACACAGTCCCTGAGCAGTCCGGACATCGGGACTTCGGCAAGATAGGGCGGGGGCTTGGTTCGAACACCTGGATTGTCAGCGTATCGGATGGAGAGCGCCTGGCCCCGGTGGGCACTGTGGGAGAGCTGTGGCTGGAAGGGCCCCTGGTGGGCCTGGGCTATCTTGATCAGGACGAAAAGACTGCCGCCAGTTTTATAGACAACCCTGCATGGCTCCTCCGAGGAGGTCCTGGTGTTCCTGGTCGTCAAGGCCGGTTGTACCGAACTGGGGATCTTGTCCGCTACTGTTTTGATGGCAGTATTCTCTTCCTAGGACGCAAGGATAACCAGGTCAAAGTCCGCGGGCAGCGTGTTGAATTGCAAGAAGTGGAGCATCATCTCCAGGCTCATCTGTCGGGAACAACCGGAGTCGTGGCCGACGTTGTCAAGCCCCAGGGCAGCAGCAATGCCATGCTCGTCGCCTATCTCGCGGTTGGAGAGACTATCCATTCCCCTCTAGACAGCGTCCATGCAGCATTGCGGCCCCTCACCCAAGGCCTCACGGACTCCCTCTCTGCCCGAATCCCGCAGTATATGGTCCCTAGCATGTACCTCCCCGTTGCAGAGATTCCAGTCACTACGACGGGAAAGATCGATCGAAAGCAGCTGCATGAATTAGGATCATCGCTTACAATGGAACAGTTAGCTCAAATCCAACCCCCCCAGGAAGGCGAGCAGCAGGCTCCCCAGACAGACCTCGAGAAGCTCCTCCAGCAGCTCTGGGCAGAGGTCCTCAATATCAGTCAAGAACTCATTGGGATGCACGATAGTTTCTTTTCCCTCGGGGCGACTCAATCACAGCGATGCAGCTATCGGCCAAGCTGCGCTCAGCAGGCTTCCGTATCGCCGTTCCTGATATCTTCAAGCTGAAGACAATCTCCCGACTGGCTCCAAGTGCAGCATCTGTGCAGGGGCGGATGAAAACCACCTGGGAGACCCGAGAAGACGAGCCATTCGAACTTGCTCCCGTTCAGCAGATGTTCGTCAACGTGGTGCGCCGCAAGTGCAATCATTTTAACCAGAGTTTCTTCCTGCGCATCACCCGTCAGGTGCGGGCTGAGGATGTGCGGCGAGCCCTGGACTTGATTGTGACTCAACATCCTATGCTGAGGGCACGGTTTGCCTCGGATCAGGGCGGACATTGGACGCAGCATATCAAGCCCTACACCCCCGGTTGCTACCGGTACTGCGAACACGTCGAGTCCTCTCTGGCCGAGGCCTCGCCGCTGTTGGATGCCAGCCAAGTCGCCCTTGATCTGGAATCTGGACCGGTCTTCTCTGGGGATCTGCTCCAGATTCGCGACAGCGGTGATCAATATCTGTACCTGGTGGCGCATCACTTGGCGGTCGATCTGGTATCCTGGAGAATCATTCTCGCGGACATGGAGGACCACCTGACAGCAAAAGCCTCTTCGTCTTTCACGCCGATGCCCTTCCAGGCATGGTGTCAACTTCAGGCCGACCATGCTCGAGACTGCCTCACTCCCGAAACGGCGTTGCCCGTTGAGATTCCTCCTCCGCAGCAGGGTTACTGGGGAATTGATCCTGAACGCAACACCTTTGGCAACGCCATCCACCGCAGCTTCACAGTGAGCAAGAACGTCACAGATATCCTCTTGGGGCCTGCCAATAAGTCATTTGATACCCAGCCGGTCGAGATTCTACAGGGTGCTTTGATATACTCCTTCACTCAAGTGTTTCGGGATAGGGCGCCTCCTACAGTGTTCAGCGAAGGGCATGGGCGCGAGCCATGGGATGCCGCGATCGACCTATCAAGAACAGTTGGCTGGTTCACTACTATGCTTCCGACCGTAACCTCGATGGGACCGGATGATACTCTACCTGATATCATCCGCCATACGAAAGATAACCGACGTCTAGTGCCAGGAAACGGATGGCCATACTTTGCATCCCGCTACTTGAACCCTGCTGGAAAGAAGGCATTTGGGGCCTACGGTCTGCCCGAGATCACATTCAACTACTTGGGACTCTACCAGCAGTTTGAAAAAGACAGTTCCCTGCTTCAGCCAACCGCCTTGCCCTCAGGCAGCATGTCGGATGTTGACGAGCGGATGCAGCGCTTTGCCCTGATTGAGGTTTCTGCATCGGTGTCCCAGGGCTGTCTGGAGTTTAGCTTCTTGTTCAGTCGTCATACAAAGCACTTGCGTAACCTTGAGGACTGGGTGGCCCACTGTCAGCATGCTCTACGAACGGCTTCGGAAGTCTTGATGCAGCGCCAGCCAAGCTATACACGATGTGATTTCCCACTTATGCGCCTGACAGAGCCAGCCTTGCAGACTCTCGTTGGATATAGTCTTCCACAGCTCGGAGTCTCGTATGGTCAAGTTGAGGATATCTACCCTTGCTCGCCGATTCAGCACGGTATCTTGCTGAGCCAAGCGAAGAACCCTGGCGTGTATTGGACGCGCATTCGATGGAAGGCCCGGTCCTCCGAGTCTGCATTGCCTCCCAATCCTGAACGACTAGCCCGTGCCTGGATGAGGGTCGTTGCACGACATCCCGTCCTTCGGACCCGGTTCATCGATGGCCTTTCTCCCAACAGCGTGCAAGATCAGCTTGTCTTGAAAGCCAGCGAGCCCGAGATACATGTGATATCTGGTCCGGAGGCGACAGATCCCATCGCGGCTCTTGACTCTTACTGGGAGCAAAGACAGCGGAAAGGACACAAGCTTCACTCTTTCGTCCTCAGCCAGGCACCCTCCGGCGAGGTCTTCTGTGACCTCGAAATGAACCATGCAATCACGGATGCCACTTCGACTGCTCTACTCAAGCGAGACCTTCAAGCCGCATACACTGAAACCCTCCCGACTATGCCGGGTCCTTCGTACAGCGATTACATCAGGCACATTCAAGCAATACCCGCGGAGGTGGGGATAGACTACTGGCGACGGTATCTCGACGGCGTTCATCCCTGCATTTTCCCAACACTGGCAGCGAATGGCCAGGACGTGGGGGAGAATCCAAGAGGCTCCATACATCATATGCTCGATCAAGACACGAATACCAGTCTTCAAGTTTTTCTCAAGACGAACGAGCTGACCGCCTCCAACGTCTTTTACCTCGCATGGGCCCTCATTCTCCGGTGCTTTGTTGGGTCGGAGACAGTGTGCTTTGGCTATCTAGTATCCGGTCGTGATGTCCCTATTGACAATGCAGACAGGATAATCGGGCCGTTTATTAACATGCTGGTCAGCCGAGTTAGGCTAGGAGAGGGGGTCACTCTGATGGATATCATGAAGCAAAGCCAGGCGGATTATCTGGACAGTCTCAAGCACCAGCATCGCTCTGTGGCGCAGGCCGCTCACTCTTCGGACGGGCCAGCCGAGCCTCTATTTAATAGCGTCATCTCGGTGCAAGGTATGGATCTGAAGAAGAGCAGCAAGAGTGAAGAACGTGGTCTCTGTCTTGAGGAGCAAGACGGTAACGATCCAACAGAGGTACGTATTCCCATCCCTTACACATGCATTTGTGGACAGCTTTCCCCTAACCTGGTGATTTCCAGTATGATATCATGATGAATATTGGTCTGGGGGACCAGGAAACAGCAATGACGTTCAGCTATTACAAGTCCCTGTTATCGGAGCAACAGGCAAGCGGCATAGTGGAGTCCCTGCTTCAAGCGGTCCGGGAAATCATCCACACTCCCTTCAAGAAGCCCTCCCAGGTGAATCTGTCAACGGGCCATGACCAGCAGACTATCTGGAGCTGGAACGCCTCCGTCCCGCCCACAATCAACGTTCCAGTGCACGGTCTGATTGCCAACAAGGTGAAGGAGCAGCCCGATGCCGACGCAATCTGCGCATGGGATGGCGAGTTGTCGTACGGGCAGCTAGATAGCTTATCAACATGCCTGGCCCATCATTTGGTGGCTTCTGGAGTCTGTTCGGACACGGTGGTTCCTCTCTGCTTTGAAAAATCCATGTGGATGCCGGTGGCCATGCTCGCTGTGATGAAAGCCGGTGGTGCTTCAGTCTCCATGGACGCGAGCCTGCCAGAAGAACGATTGCGCACAATCGCAAAGCAAACGGAGCCCGCGGTGATTCTCTGCTCGGAAGCAACCAGCGAGAAGGCGGGTCGACTTGGTGCCCCCCGAGTGATTACTGTCGGCCAAAGGTTGCTCTCAGGGCTGACGATCCCCAGTCGAGGCACGACGCTCCCCAAGGTCGATCCATCTGACCGATTATACGTGACTTTCACATCCGGCAGCACGGGGATTCCCAAGGGAGTCATGATCACGCATGCAAACATGAGCAGCGCCCTGGTCCAGCAGCAGGAGGCCCTCTCATTCGGGCCTCACGTCCGCGTCTTTGATTTTGTGTCCTACGCGTGGGACGTGGCTTGGTCCAACCTGCTCCGTACTTTGGTTGCTGGTGGCTGTCTCTGTGTCCCATCTGACTCCCAGCGCCGGGAAGGGATTGAAAAGGCCATGACACAATTAAGAATCAACTGTGTTACTCTCACACCCAGTGTCGCTCGTCTGCTCAACCCTGCTGCGGTGCCGCACCTGAATACCCTGGCTCTAATTGGGGAGCCCCTGTCGCAGGCCGATATCATGCGGTGGGCTCCTCACACCAACGAGATCATCAACACCTACGGGCCATCCGAGTGTCCTGCATGCGTCACGGTGAATCGCATCCCGCTGGACCTCTTGTATGAGCCCAATCTCGGGGCGGCTTCTGCCTGCAACACTTGGATTGTGGATCCCGATAACACCGATCAGCTGGTTCCAGTTGGTGGTATCGGAGAGCTCTGGCTGGAAGGGCCCCTCATTGGGCTTGGATACCTGGGCCTCCCCGAACGAACGGCCGAGAGCTTTGTCGAGAATCCCCGCTGGCTGCGCTCCGGGTATGCTGGACAGCTGGGGCGTGGCGGCCGGCTGTATCGCACCGGCGACCTGGTCCGCTACACCGATGATGGCGCGCTTGTTTATATTGGACGCAAGGACTCGCAGGTCAAGATCCGCGGCCAGCGCGTGGAGCTGGGAGAAATCGAGTACCAGATCCGCGAGGGCATCGCGCGCATCTCGTCCGCCACCGGCGATCCCACCGTCGTGGCTGGGGTTATTACTCCGCGCGGAAGCGGCAGCAAGATGCTGGTCGTCTATCTTGGCCTAGGGCAGATGGCCACAGGCCCTGTTGACCGCATTCGAGAGGCTCTGGCGGGCTATACAAGCGGGCTGGATGTGTATCTCTCGGAGAACCTCCCCCACTACATGATCCCCAATGCATACATCCCCGTCTCTGAAATTCCAATGACAGTGACCGGCAAGACGGACCGCGGACGCCTGGCCCGTACCGGGGCATCATACACCCTGACCGAACTGGCCGAGATGCAACCCTCGCGCGGCAGGCAGCGGCCCCCTACCACCGCTATGGAGCGGCGGCTCCAACAGCTTTGGGCGGCAGTCCTGGGGCTCGATGTCAGTGCCATCGCTGCCGATGACAGTTTCTTCCGCATTGGCGGGGATTCGATTGCGGCGATCCGCCTGGGCCAACGTGCGAGCGAGGATGGTCTGGCCGTGGCCGTGGCCGACATCTTCCGCAAGCCCCGTCTCTGCGATCTGGCTCTCCTCGCTCGTGAGGACCAGGTCACCTACCTTGATCCTTCTCCCTTCTCCCTTCTCCCTGCCGGTTCTACCTCGGACTACGTGACAGCTGGTCTCAGTCCGCTCCTGGACTGCCCACAGCACCACATCGCCGACGTCTATCCCACCACGGACCTCCAGAGCATCTACGCATCCTTTGCAGTGAATGCACATCGGGGGGAGATCGAGTACATCTACATGGACCTTCCTCAGGGGGTGGACATCGCACGCGTCCGCCGCAGCTGTCTGGATATGTGGCACCATCTCGACATCCTTCGCACCGTTTTCATTCTTGATCAAGGTCGTCTATTGCAAGTCGTGCTCAACAACGTGGAGCCCGAGATTACCATGCAACACACCGAAGGGGATCTGGCAGCAGCATGCGAAGAAGCCTACAGCTGTGATCTACATGAGTCGCTGAGCCTGGGCCGGTCGTTCACGCGGTTCTTCATCACCTCCACCGCAGACGGTCGAGTGCGATTCACCATCCGGCTCTCTCACGCTCAGTACGACGGCTACAGTCTGCCTGTCATCTTCTCCGTCTTTGCGGCCTTCTATGAGGGCCAGACACCACCCCCAGCCCCGAAGTTTTCGGGCTATATCCGCCACATGCAAAAGCAGCAAGAGGCCGCGTATCCCTACTGGCGCACGCTCTTGCAAGGCTCGTCCATCACCCGGACCAGACACCTCTCCATAGCCGATGGCCACTGTCTGCCAAACCAGCCCCGTGGCCGGCTGGTTCAATCCAAGACCGCTGTGCCTGCCCCGTCAGATCAACCGGGATTCACGTCGGCAACGGTCTTTACCACCCTCTGTGCTCGCATGCTGGCACAGATGACAGGCGTCAGCGACGTGGTCTTTGGGACCATTGTCTCCGGCCGTTCCTCCTTACCAATCGCCCTCCAAAACGTCGCCGGGCCATGTGTCAACACCATCCCCGTCCGCGTGCAGATCGAACCCGACCAGCCACTAGAGGAGCAGCTGGCCAGCGTCCACGACCAACACATCCAGGGCATGTCTTTTGAGACGAGCCAGTTCAGCGACATCGCCGCCCACAGCACCACTTGGCCCGAGGACTCGCGCGCTCCCGACCTGGTTGTTCAGTTCCAGAACCTCGACAACCTGGAGCATGACAAGGGTACGGAAATCCAGGGGATAGGCAGTACACTTGCACCGTGGGAGCAGCGGAACCGGTCCGTGGACAGCGATTTCCTCTTTATTCTGGCGAAGCCGGTGGAAGAAGCCTGGGAGATATCGGTGTCGGCGAGCTCTAAGGTATACACCCAGGATACTCTGGATGCCATGCTCAAGGCCTTGTGTTTGCATGTTGAGAATGTTTAA | MPTENGEVQHRIDGLQDQDTVEILLVTLESWALQAVEEFKASKGMLEIGSVSDRSNDARKVKIELDGGTENILAQLRGEWVRQLVHQHHEAASNFAASIILDGSPERLLVLRFMERREAWVSEEAPKSWCQERRCAPFLEMHCQARPEELQFHIVSPPPVAAQAFVQHLDLRIRQSISNRSLADVPVETVPAASDSELRAIWGWNGTLPEGVQRCVHDIIAQQARLRPHAPAVSAWDGELDYQQLDRISTRLAHYLVQLGAGPDDIIPLCFEKSKWMMVAILAVMKSGAVIAALDPTQPEDRLRSIVKQLQPRWILTSPAQIEVAARLEISNVIALDEGRLRQLPDSEGRHLPCVDPSRNLYIVFTSGSTGTPKGVMINHTNFSSAIAYQHEALGMDNTARVFDFASYAFDLAWGNIIHTLAAGGCLCIPSESERRGNIAEAIRRLGVNHLQLTPSVARLIDPRDIPAVRWILLIGEPMTQADVAQWTPYCKLINSYGPAECTVAVTFQTIPHGRPWDSSMGKGVACSTWIVDEEHGETLVPLGHTGELWLEGPLVGQGYLGDLEKSAASFIDSPAWLTRGIPGVVPGRRGRLYKTRDLVRYNPDGSLVYVARKDTQIKIRGQRVELGDVEYHLKLALPDKIPSVAAEAITPRGSSSTILVAYLALGEEATGAAESTRESLASCLHGVEEYLADRLPRYMVPSLYLAVPEIPMTATGKTDRLRLREIGSSLTLDQLAALQPSRAVEIKAPETEMEHRLQQLWAATLNISPSSIGTGDSFLRIGGESMAAIRLVQLARKEGIILTVADIFNHPRLGDMAQEAQRERAQDVTTIPPFSLLRRGHDVRDACALAAAECRVSANSIADLLPCTPLQEGLLALTVKQAGEYVRQMVSELPVDVDLARFRAAWAQVIHEAAILRTRIVDLPNAGLLQVVVADQPEWATGSDLGHFLESEKGKPMGLGTSLARFGLVSDQAQGKVLFVWTIHHALYDGWSLPAMLERVEAIYAGGSCDMLPSFAGFVKYLADGTVEDAQGYWQSQFNGIQAAVFPALPSPDYQPKCQDLLQYHVASVSWPGNDITASTAVRTAWAIVASRYTQSPDVIFGATVSGRQAPVPFVERMAGPTIATVPVRVNVQGDATVASLMQSVQTQAVAMIPYEQTGLNQIRRINSDADQATQFQSLLVVQPPSSKTSRRPDECLFRVDLDAGDEFRSINTYSLMLECHLESDGMRLRMSYDRLVIDAEQVKRIARQFEGVLRQVCSEVNAQELVGTVSAASEADLAQVWAWNATVPQTVQGSVNDLIAQRVQQQPDALAICAWDGQLSYQELDTLSTQLALSLVQQGAGRGSVIPLCFEKTMWTPVAMLAGMKAGSTVVTMDPSQPEDRLQSIIQQTQPPLILTSDTYKPLASRLAGVVIRVNRDTLQALPQHDLDTPDLPTIHPTDGLYIAFTSGSTGNPKGAVMSHQNIRSAIHHSTGALGFTPTTRVLGFSSYAFDAVWVEFLYAMAVGGCLCIPSDAQRNSGDLAGCMEGLRVDLALLTPSTARLLDAEAVPSLKTLVLIGEPVTSDDLARWTGNVDLRNGYGPAECSAITAAYKFQGPNDQPSIVGHAVGLVAWVVEPSDGASLSPLGAVGELWVEGPLVGKGYHGDPEKTAMSFVHDPVWLLQGTLGYPGRPGRLYKTGDLVRYTSDGKLVCVGRKDTQVKIRGQRVELAEIEHYIKEATRASVVVDMARPQGSRGPVLVAYVALGQVAAVGSPEAVRTALRRCIQGVEDHLSKHLPRYMVPSFYIPVVDIPLTATGKTDRRALRNTGSSFAIDQLAELQPSGGQKRTLPQSEMERSLQQLWAEVLNINPSQIGMDESFFLLGGDSISAMQVSAKSRARGVRVTVPDIFKFKTIARLARREIQIDDLAIDDREMLDTPFALSPIQQFFFDAQKDKQGHFNQSFLVRTTKPQQPDAVLRAVQFIVARHSMLRARFHRGTDGVWTQQITSRADESYACRHHRLASVEDAVPVLNSSQRSLDVQAGPIFAVDLIDTPPDCQYLFLTAHHLVVDLVSWRIILEELEEYLLTGTVAGSSAPLSFQTWCRLQERHAQDHLTPGTAFPFDIQPPCEAYWGLSPELNTHDAIQEAGFTLSKKTTSILLGPANHAFQTQPIELFLAALLHSFMDSFTDRDPPTIFNEGHGREPWNSAIDLSRTVGWFTTLAPISVEAGSHHLSIPQFVRYIKDRRRQIPHNGWSYFTSRYLNTEGKTVFSRHACPEVTFNYFGLYQQLERSDALFQSCPSLQDRVLDVADRMARFALIDVSAEVTQDCLQFRFLSSQRMQKQDALARWIAACERLLEATAVQLLDMQPSYTLADFPLLPPTDATWKQLETLPRLGLAYGDVEDIYPCSPLQQGILLSQAKSPEMYWTRVRWVVRSSTGASSPVDADRLERAWRKVVARHAALRTRFSDSPFSDGYSYQIVLKDPPPTIHTIQAADPMDAVVEYRAASGSQVRPVHSLVLCPMASGDLVCDLQINHAIIDAISIRVLKRELCAAYDDALPADPGPRYSDYISHLQSLPTTEAKEYWQTQLAGAQPCIFPTLNEPIAEPKDAAAVTSLPISPETDQALRQFCRSHALTPANVFSLAWMLVLRSYTSSESVCYGYLISGRDVPIHQVDRAVGPFINMVVNHVEIDSNRTLLAMLQEVQAGYLGSLKYNQYPLAEILHDLSTEGQPFFNTVLSVQSGGRADSAQPPDTPTITLENETWHDPNEYDIAASVLLLDDNPRITLNYSRHLLSERQASAVGATFVEALANIVRHPDRNLCDLETTISPQDLATIWDWNAKVPATLSSSVPELFSRWVKQQPDAQAVCAWDGDLTYRQLDEGSSRLAHHLLARGVRPQSILPLCFEKSRWVPVAMLGVLKAGCAVVTMDPEQPVERLQLVVQKTQGVILTSPACQGLASQLRPEAVVVDGRSLERMPQPDPLSLPTIHPTDRLFLVFSSGTTGTPKGSVMSHQNACSAVHHQQARLGLPPAARMIDCLSYAFDAPWFNFLHAFSSGGCLCVPSDQQRKDNLAGCIESLKANYALLTPSIARVLDPAAVPGLETLAFGGEAIRAEDISRWHEVKLMGYYGPSECTVVSTIHQFESKTEEPRMLGYSSGLTAWVVDPLQGRRLAPLGATGELCVEGPLVGQGYIDEPQKTAASFVEDPSWLLRGGGPDFPGRRGRVYKTGDLVRFRPDGSLLYVGRKDTQVKIRGQRVELGEVEHHIRQTLPAGVDVPVVVELVHPHASANPVLVAFLPVGKEAAGSSESMRAVLDRYTHGVEDRLMARLPSYMVPRMYIPVPAIPFTTTGKTDRRSLQRVASSYTLEQLAALQPSRDTRRAPTTAMERRLQGMWATILDIDAVTIAATDSFLRIGGDSIGAIRLVRLAAEQDILLTVAAIFKSPILCDMAQVATLGSTSGPHDIPPFSLLKQDVDISQARLQVAAWCELSPSAVEDILPCTPLQEGLLSLTVKNQGAYVNRQVLQLRDEVDLGRFQTAWNQVAMATSILRTRIVDLPGQGLVQVVTSEVPRWNHGRSLDALVQDDLQRPVTLGTPLATFGLVDAGDVGGDRQRYFLLTLHHALYDGWSLSLLLDEVSKAYHGTASDGLVSFKTMIKYVTELGAEADSYWQGSLDGLVAEPFPALPSPMYQPRAQDILEHNVLGLRWLQNHHITPATTLRAAWAILTTHYTQSADVLFGSTVTGRQAPIHRVGLVEGPTIATVPVRIAIQEKATLAGLLDQVQEQSVDMIPYEQVGLQRIRRLSADTERACQFQTLLVVQPAPEPSATDGPSALFHAEDDTLSQAALSSFNSYALLLQCQLTATGVSMQMSYDSHVMAQPQVQRLARQFEHLVRLLCDESQHSSSVAHIDAVCEDDLRRIWALNSPPPARVDACMHDLITEQAQRRPVSQAVAAWDGELTYRELDELSTQLAYGLVNLGVGPHAVVALCFEKSMWMPVAMLGVMKAGGASVAMDITQPEDRLRTVVQQVQPPLVLSSLEAEDLARRLSDRPVHVVSQDRLQTTSLGSGQGGQLPPVQPTDRLYIAFTSGSTGVPKGAVMSHQNFTSAVHHQTALEFTASARVFDFSSYAFDACWLNFLHTMAAGACLCIPSEEERKSDITGCMRRLAVTYANLTPSTARLIDPTSVPSLQTLVLVGEPVAQQDITQWKAHVQLKNGYGPAECSAISTTFDLGQSDHDPATIGAGRGMVTWVVEPTESRHLSPYGAVGELWVEGPLVGLGYLGRPDLSAASFVDNPPWLLRGGPQGFPGRHGRLYRTGDLVRYNLDGTLVCIGRKDAQVKIRGQRVELAEVEHHLHQALPSAAVDVSVAVEVITLQGSANPLLVAFVAMGEAALGPRETVRAKLALYSQGARERMADQLPVYMVPSLLLPAVEIPTTATGKRDRRRLREMWASKSLEELAELQPTKGNHQAPTTDVERRILQLWAECLNISASKISIHDSFFALGGDSISAMQLSAKGRSIDLQMTVGDIFKYKTIARLALSISPAVESAVLHAPEGQGASFALSPIQQMFADMQQGVSNHFNQSFFVQVRRPVTLPQIQAAVDALVAHHGMLRARFKCSRENIWSQWILPPGTTGSYRVCQHEVAGLQAASAVINHSQESLDIQNGRLMAVDLINADEGQYLFLVVHHMVVDLVSWRIILADLEEHLTTASLSGFTSMSFQTWCQLQADHHAESPFELEAVLPNGAPPPLLPQLDYWGQVRSLNTFDNMIKDGVVLGRQDTETLLGPANAAFDTQMVELLHAAILHSFANVFHDRAPPTVFSEGHGREPWRSAIDISRTVGWFTTMFPVVATAKKGDSIASIVRHVKDRRRQIPDNGRPYFASRFLTPAGKRAFQVNGPVEVIFNYLGLYQQLERSDALFHMREMPDGVDDMADISGSLFRFALVDISASVTDGCLHVDFMYNRHMQHQASIRAWIKECRQSLHAAAQELPLIQPSYTLCSFPLLRMTDPALAILQQRLTELDLAYGQVEDIYPCSPLQDGILMSQIKNPDLYRTRIRWMAQSAQGSTAVDTNRLKHAWQQVVDRHPALRTIFVDSISGRGLKDQVVVRNLRADVHIVQSADEASIEDMSPTKGKTSSVLTLSTTNSGVLCELSISHALIDAFTLGILKQELCAAYTGLLSSSLAPLYSDYIHFIQSLPAGSAAAYWQEHLQEVKPCLFPSLGGLNTEGRRSQAHIPITFERELHLALRIFCMEHGLTVSNIFHVAWALVLRAYTGLDTVCFGYLTSGRDIPLQGADGTVGPFINMLASRVDLGSKDSLMTLVQRDQEQYLNSLEFQHYPLAKIFHLIDTPEKGLFNTAMSVQASDSGPKDCQSAISLVDEGGDDPTEYDIMINIGVGDEDTGCNFTFNESVISDRYAKSAIDLFLHAVSHIVQHADQTAQEANFISKQDLQSIWQWNAAVATPLDWCVHELIIEQAEKQPAAAAICAWDGNLTYKQLNDLSTQLACHIRQLGVGPGVNVPLLFEKSRWMSVASLAVMKAGGTMVGLDPGQPAGRLRSIIDQVQPCLILTSAGNSNSAAAFASCHVVRVDDASLAQLATPCNALLPPVDPASSLYLVFTSGSTGVPKGVAISHSNLSTAITHQKRILKLSVASRVLEFASYAFDVSWGTILHTLAAGGCVCVPEESERRGDISAAMRRMEVNYAHLTPSVARLLNPSNVPLLQTLVLSGEAVSRADMEQWSRHVHLINAYGPAEAAVWVIFAHLDSALSMPSIGKGGGCTTWIVDPSRPDQLAPVGCTGELWLEGPLVGRGYLHDPERTAAAFIENPRWLLQGAGGAHGPGRQGRLYRTGDLARYTPDGSIVYIGRKDNQIKIHGQRVELEEVEKYIEQAMLNSAAAPAVPVVAAVVTPQGSKKAILAAYLALGEPATGSVEIVRKSLSRYTGIINPALEESLPTYMRPSIYIPVAEIPTTTNGKADRSKLATMASSRTLAEWAGLQTSEGKWRPVSSPGELGLQRLFAEVLNMDQSLVGMDDSFFSLGGDSITAMQLSAKSQSSLLYITVGDIFKHKTVAQLASNARQTASTSVQLPEVPNSLFELSPIQQLFFASQDKGKNLFNQSFLVRVSRSLNPNELQKAIGVLAARHSMLRARFVQSADGVWQQKIVNDSAGCYTFRSHHITTLQDMEPLLHSSQELLDIVQGPILAVELIDSSHDGQYLFMTAHHLVVDLVSWRILLGDLEEFIRSGTITGFPPFSFQSWSQLQANYARDHLPPKIALPFKVSPPRHEYWGLVPGHDANTLSDSGRGSFTMNKRLTDILMGTANSAFDTQPVEILHAALLYAFAQTFQDREAPPLFTEGHGREAWDSGIDLSRTVGWFTTIFPVAASITQNHSLPEVVRRVKDTRRQTPRNGWSYFTSRYLNTDGRQAFQIKGPVEIIFNYMGLYQQLERPDSLFQQCDIAVTAPPAAAATLSRFALIDVAASVVQGQLKFEFLYNREMKGQDRILEWISKTESSLKAAAEELLLQSPSYTICDFPLLSLTDQGLDELLNRVLPAVGLSYGQIEDIYPCAPIQEGIIMSQAKSPELYWTRVRWTVQSTGTSPVDLDRLRRAWQMVVNRHSILRTIFIDGIGSGQVKDQVVLKDFPVDVEVLHAELHQLAEPVTGRQWHTDLSSRRDTPQHSLVLTQTASGVVFCDLELNHAMVDAYSLALLRQEICAAYTGSLPATPAPAYQAYIEHLQGLSLVEGQRFWQTYLDNAQPCHFPALGQPDVTDGSNARRALSISLDAATHQALRVFCQQHAVTPSNVFYLAWGLLLRAYTGLDTVCFGYLTSGRDVPVPGVDRIVGPLINMLVCVLEFRTGASVRSAMQKVQEDYLAALQYQSTPLSKILQLSGTSGRGLFNTGISVQGGATSGEPDEHDIIVTDQTGLDSPEYDIAVAISHDEEETEIAFDYMAVTLSDQGGQSLAGLFVQVVADVIRAPDQTVQAVNVISKEDLQSLWTWNRTVPETVQACVHDMIGQKACDSPDSPAIHAWDGALTYQELDLLSTRLARYLIDLGVRQNTAIPLCFEKSMWMPVAALAVMKTGAACVAMDMTQPEKRLQAIVDLVQPDLLVTSVANRKTVQQLADTKAVLAIDQAFFSQIAPPTSSCLPTVSPSSPLYTVFTSGSTGTPKGAVISHANFASAIVHQTDLLALDPDSRVFDFVSYAFDVSWSNLLHTLAAGACLCIPSEAMRRDNPVEAMQAMKVTHAQFTPSMARTVDPDQCKTLKALILGGEAMSQHDIAVWAPWVDLRVAYGPAECTVAAVMDTVPEQSGHRDFGKIGRGLGSNTWIVSVSDGERLAPVGTVGELWLEGPLVGLGYLDQDEKTAASFIDNPAWLLRGGPGVPGRQGRLYRTGDLVRYCFDGSILFLGRKDNQVKVRGQRVELQEVEHHLQAHLSGTTGVVADVVKPQGSSNAMLVAYLAVGETIHSPLDSVHAALRPLTQGLTDSLSARIPQYMVPSMYLPVAEIPVTTTGKIDRKQLHELGSSLTMEQLAQIQPPQEGEQQAPQTDLEKLLQQLWAELSAKLRSAGFRIAVPDIFKLKTISRLAPSAASVQGRMKTTWETREDEPFELAPVQQMFVNVVRRKCNHFNQSFFLRITRQVRAEDVRRALDLIVTQHPMLRARFASDQGGHWTQHIKPYTPGCYRYCEHVESSLAEASPLLDASQVALDLESGPVFSGDLLQIRDSGDQYLYLVAHHLAVDLVSWRIILADMEDHLTAKASSSFTPMPFQAWCQLQADHARDCLTPETALPVEIPPPQQGYWGIDPERNTFGNAIHRSFTVSKNVTDILLGPANKSFDTQPVEILQGALIYSFTQVFRDRAPPTVFSEGHGREPWDAAIDLSRTVGWFTTMLPTVTSMGPDDTLPDIIRHTKDNRRLVPGNGWPYFASRYLNPAGKKAFGAYGLPEITFNYLGLYQQFEKDSSLLQPTALPSGSMSDVDERMQRFALIEVSASVSQGCLEFSFLFSRHTKHLRNLEDWVAHCQHALRTASEVLMQRQPSYTRCDFPLMRLTEPALQTLVGYSLPQLGVSYGQVEDIYPCSPIQHGILLSQAKNPGVYWTRIRWKARSSESALPPNPERLARAWMRVVARHPVLRTRFIDGLSPNSVQDQLVLKASEPEIHVISGPEATDPIAALDSYWEQRQRKGHKLHSFVLSQAPSGEVFCDLEMNHAITDATSTALLKRDLQAAYTETLPTMPGPSYSDYIRHIQAIPAEVGIDYWRRYLDGVHPCIFPTLAANGQDVGENPRGSIHHMLDQDTNTSLQVFLKTNELTASNVFYLAWALILRCFVGSETVCFGYLVSGRDVPIDNADRIIGPFINMLVSRVRLGEGVTLMDIMKQSQADYLDSLKHQHRSVAQAAHSSDGPAEPLFNSVISVQGMDLKKSSKSEERGLCLEEQDGNDPTEYDIMMNIGLGDQETAMTFSYYKSLLSEQQASGIVESLLQAVREIIHTPFKKPSQVNLSTGHDQQTIWSWNASVPPTINVPVHGLIANKVKEQPDADAICAWDGELSYGQLDSLSTCLAHHLVASGVCSDTVVPLCFEKSMWMPVAMLAVMKAGGASVSMDASLPEERLRTIAKQTEPAVILCSEATSEKAGRLGAPRVITVGQRLLSGLTIPSRGTTLPKVDPSDRLYVTFTSGSTGIPKGVMITHANMSSALVQQQEALSFGPHVRVFDFVSYAWDVAWSNLLRTLVAGGCLCVPSDSQRREGIEKAMTQLRINCVTLTPSVARLLNPAAVPHLNTLALIGEPLSQADIMRWAPHTNEIINTYGPSECPACVTVNRIPLDLLYEPNLGAASACNTWIVDPDNTDQLVPVGGIGELWLEGPLIGLGYLGLPERTAESFVENPRWLRSGYAGQLGRGGRLYRTGDLVRYTDDGALVYIGRKDSQVKIRGQRVELGEIEYQIREGIARISSATGDPTVVAGVITPRGSGSKMLVVYLGLGQMATGPVDRIREALAGYTSGLDVYLSENLPHYMIPNAYIPVSEIPMTVTGKTDRGRLARTGASYTLTELAEMQPSRGRQRPPTTAMERRLQQLWAAVLGLDVSAIAADDSFFRIGGDSIAAIRLGQRASEDGLAVAVADIFRKPRLCDLALLAREDQVTYLDPSPFSLLPAGSTSDYVTAGLSPLLDCPQHHIADVYPTTDLQSIYASFAVNAHRGEIEYIYMDLPQGVDIARVRRSCLDMWHHLDILRTVFILDQGRLLQVVLNNVEPEITMQHTEGDLAAACEEAYSCDLHESLSLGRSFTRFFITSTADGRVRFTIRLSHAQYDGYSLPVIFSVFAAFYEGQTPPPAPKFSGYIRHMQKQQEAAYPYWRTLLQGSSITRTRHLSIADGHCLPNQPRGRLVQSKTAVPAPSDQPGFTSATVFTTLCARMLAQMTGVSDVVFGTIVSGRSSLPIALQNVAGPCVNTIPVRVQIEPDQPLEEQLASVHDQHIQGMSFETSQFSDIAAHSTTWPEDSRAPDLVVQFQNLDNLEHDKGTEIQGIGSTLAPWEQRNRSVDSDFLFILAKPVEEAWEISVSASSKVYTQDTLDAMLKALCLHVENV |
| ungB’’ | ATGTACCCAGCCAAGTGGGTGTTCCTGATCGGGGTGGCGATATTCGAGATCGGATCTCTGGTCTGCGGAGTGTCTCCCAGCTCGGGTGCCCTGATTGCGGGCCGCAGCATCTCAGGACTGGGCGCTGGCAGCATCAACGCCGGCGCTGTCGTGATCGTGACCAACACCATTCCTGTGCGAAAGCGGCCGATCTACCTGGGCTGTCTGGGGTGTGTCCATGGGCTAGTCAGCGTGTCGGGGCCATTGTGAGCTCATCTCCGTGCAACCATAATCCACATAGATCGGGGGCTAATGCCGTGTGACGAACAGGATTGGGGGTTTGCTGACGGACCATGCCTCGTGGCGATGGTGTTTCTATCTCAATCTGCCCATCGGAGCAATCACCATCCTGGGCATCATTTTCTTTCTGTCGGCTAACCAACCCCCGTCGGGACATCTCAGCTGGAGGGAAAAACTTCAGTCCATGGATCTTCTGGGCTCGGCGTTCTTTATTCCTGGGATCGTCACGCTTCTGCTAGCCCTGGAATGGGGAGGGTCACAGTATGCCTGGGACAATTGGAGGATCATCCTTCTCTTCATCCTGTCCGCCATCTTGCTGATCGTCTTTGCCGGCGTTCAGATTCGCGCCCCTCAGGGTAAAGCGACACTCCCCCCTCGCCTGGTTTCAAATCGGAATATCCTGGGTCTTATCGGCTATATTGTTGGTAACAGTGGGGGGTTGTTTGTCTTTGTGTACTATGTATGTTCTCGACCTGGATCCCCCCCCCCCCCCCCTTTTCCCTAGCTGGAATCTCCGCAACATGGCTTAACGACTTCGTCACCTTGACATAGCTGCCAATTTGGCTCCAAGCCATCAAGGGGTTCTCGGCCTCCAAGTCAGGTCTTTCAATCTTACCCACCCAGTTAGGGGTGGTGATTGCATCTCTAGCTGGCGGGGTCTTAGTCACTTTCGTCGGATACTACACGCCCTTTCTGATTATATCCTCCCTAATGGCCGTCGCTGGCGCCGGCATCCTGAGCTCCCTGCATCCGTCCTCCAGCCTCGCAAGCATCCTGGGGTACCAAGTGGTGCTGAGTCTGGGCATCGGGCTTGGATCGCAGAACGCCATGGTCGTGCCGTCCGTGGTTTGCGCCCCGGACGATGTAGTCACGGCCATTGCGATGCTCTGCTTCCTGCAGATGCTGTCCAGCTCGGTCGCTCTGTCCATTGGACAGACAGTTTTCCACAACCGGCTGGTGGCCAATTTGCATCGCAGCGCCCCGTCCGTGGACCCCTCCCTCGTTGAAGAAGGGGCGACGCTGCTCCGCGATCGCGTTCCATCGGAGCTTTTGCCGTCTGTTCTTGGGGCGTATAGCAAGGCTGTCTCGCAGACCTTTTACGTCGGGGTGGCCATGTGTGCGCTGTCGCTTCTTGGGTCGGCCTCCATGCAGTGGAAGCGAGTCCCTGGCAAGAAAGATGCGGCAGAGGAGGAGAAGGAGGAAGAGTCGCGCGACAGTGATACCCCTGCACCACCTGCGACAGAAGAGAAGACGTGTAAGCAATAGAAGGGCGTGGGTGTATACCTTGCATCCCCCATGTGTGTTCTGA | MYPAKWVFLIGVAIFEIGSLVCGVSPSSGALIAGRSISGLGAGSINAGAVVIVTNTIPVRKRPIYLGCLGIGGLLTDHASWRWCFYLNLPIGAITILGIIFFLSANQPPSGHLSWREKLQSMDLLGSAFFIPGIVTLLLALEWGGSQYAWDNWRIILLFILSAILLIVFAGVQIRAPQGKATLPPRLVSNRNILGLIGYIVGNSGGLFVFVYYLPIWLQAIKGFSASKSGLSILPTQLGVVIASLAGGVLVTFVGYYTPFLIISSLMAVAGAGILSSLHPSSSLASILGYQVVLSLGIGLGSQNAMVVPSVVCAPDDVVTAIAMLCFLQMLSSSVALSIGQTVFHNRLVANLHRSAPSVDPSLVEEGATLLRDRVPSELLPSVLGAYSKAVSQTFYVGVAMCALSLLGSASMQWKRVPGKKDAAEEEKEEESRDSDTPAPPATEEKTCKQ |
| ungC’’ | ATGACAGTGACAGTACCGCAGGATCTCCGGCTGGGAGAGCCATATGGTGATCGTGCAAACAGTCTGTTGGTCCAAATCCCAACATGGGCGGACATGGTCGATTTTGGGCAGGGAAAGGCGCAGATGAACCACGGCTATCCACGGTCCATTGTGCACCCTGATATCCGTCTGGTAAGGCCCTTGGCTTTCTTTGATCTCGTTCAGCTGGGGCCACCCGGATTGGCCCTGATACTGACGATGATGATGTGTTTGTCCTTCAAGCTCAGTTCCGCAATTCTGGCCAAGGTCCGCGACGACAAAGAATGCACCGATGCTTCCCTTCTATTGTTCTCTGGTGCCCGGACGGCGCTTTCCTGCAAACAGTATATCCTCCATCTCGCGCAGGCCACGGATCGGCCAACCCTCGCCGATTCGATCCGGGTGTACGTGGTGGACTTCAGTGGGTTCTTGTGTCGTGACCATCACGCTGCTCACTCGTCTGGATCAACCCTTTATGCCGTCCTCTACCCCAGTGCGGCTTGTCGGGATGCTCATGCGTTTTGGCAACGCGCGGGGCCAGGGATATCATCAAGGTTGGCACAGCGATGCCTGCGGTCCAGTTCTCGACCTCGGATCCACCCATTCACTGCTCCTACCACCGCCACCGCCCCTCCCGACACCCATTCTGTGTACGAGAAACTCCGTTGTAGAATTGCTGGCCATGTGAGACACTGGTCGAGTGATCCTCGGATCACCAAGCAGGTCGCCCCCTCCGATGTCTTCTTATATGCTTCCGGCATGGCAGCCATCTACCACGTGCATCAGTCCATTCTGGCCTGGCGAACCGGGGAGTCCGTGAATGCCGGCCTCCTGTATGGGCCGACCCTCAGTATCCTGCAAGCGGAAGGTCAGGGGTTAAGGTCCTACAACCTAGGAACGGAGGCCGACTTGAATCATCTTGCTGCCCAGCTGGAGCTGGGATCCGGGGAAGCTCGCGCGGTGCAAGCAATCTGGTGTGAATGTCCCAGTAACCTCACACTGCAGACGGTCAATTTGCAGCGGATACGTCGGCTGGCGGACCAACATGGCATTCCTGTAGTAGTGGACGACAGCATTGCCAGCTTCGCCAACTTGGATCTGTTGGGAGTGGCGGACATTATCGTCTCCTCGCTGTCCAAATACTTCAGCGGATATGCCGACGTAATGGCAGGAAGGCAAGTCAAGTCAAGTCAAGTCAAAGACCCAGGCGCCAACCTACAGACACGGTGACTGATACTCTCCTCTGCCGGCATACAGCGTCATTCTGAATCCGAATTCTCCACACTACGCCGCCCTGCACAAGCAGATATCCGCCACATACGAGAACAACCTCTTCGTCGAAGATGCGATCCGACTAGAATCCAACAGCCGGAACTTCCTCCCTCGCATGGATCGAGTCAACAAGACAACGCAACACCTGGTCACCCAGCTTTTGCCTCTCGTCTCGGATCCCGCGAGCCCATTGACCCGCATTTTCCACCCGTCCGTGTGCGCATCCCGTCCGAACTATGAGCGTCAGATGCGTGCTCCCTCGTCTGAGATGCCTCGCCCGGGGTATGGAGGCGTCTTCACGATGGAGTTCGCCGACATCGCAAGCGCGTCCGTCTTCTTTGACCACTTACACGTCTGCAAGGGCCTCTCTTTTGGCGCGGATGTCTGCATCGCGTCGCCGTATATGCAGATGACGGGGCAGGCCGGGAAGAAGCAGGCGTTGGTCAATGGCACGAATGATACTATCATCCGGTTTGCCGTGGGGTTGGAGGAGCCTGATGAGGTCCTGCACCGCATCAACACCGCGCTAGACGCTGCGACGCTGGTGTATAGGGCTAAGGGCCATGTGGCCCCGTAG | MTVTVPQDLRLGEPYGDRANSLLVQIPTWADMVDFGQGKAQMNHGYPRSIVHPDIRLVRPLAFFDLVQLGPPGLALILTMMMCLSFKLSSAILAKVRDDKECTDASLLLFSGARTALSCKQYILHLAQATDRPTLADSIRVYVVDFSGFLCRDHHAAHSSGSTLYAVLYPSAACRDAHAFWQRAGPGISSRLAQRCLRSSSRPRIHPFTAPTTATAPPDTHSVYEKLRCRIAGHVRHWSSDPRITKQ  VAPSDVFLYASGMAAIYHVHQSILAWRTGESVNAGLLYGPTLSILQAEGQGLRSYNLGTEADLNHLAAQLELGSGEARAVQAIWCECPSNLTLQTVNLQRIRRLADQHGIPVVVDDSIASFANLDLLGVADIIVSSLSKYFSGYADISATYENNLFVEDAIRLESNSRNFLPRMDRVNKTTQHLVTQLLPLVSDPASPLTRIFHPSVCASRPNYERQMRAPSSEMPRPGYGGVFTMEFADIASASVFFDHLHVCKGLSFGADVCIASPYM  QMTGQAGKKQALVNGTNDTIIRFAVGLEEPDEVLHRINTALDAATLVYRAKGHVAP |
| ungD’’ | ATGGCCATGCTGGTCGAGGACAATGACAAGTACCCCGGAGTGCAATGGACTACGCCCGTGAATGAGTTGATCCGCGAGGACTTCGTTCTGGACAACCACTGGACGACGGCAAACGTCACCATTGAGGATATCCTGTCCCATCGGACCGGAATGCCGGGCCATAACTTCTCGCTTGGAGCTGTGTATGCTGACAAGCAGGCCACTGTTCAAGATGTGGTGAGGAGCTTGCGCTTCCTCCCGTCTACCGCACCCCCACGCACGACCTACCAGTACAATAACGCCATGTATATCGTGGCGTCGCATCTTATCCAAACAGTCATGGGCGATGACCTCGGCTCTATCTTTCAAAAGCACATCTGGGACCCTCTCGGGATGAGCAGCACGTATTTCCGCCTGGATGATGCACTTGCCAGTCAAAAGCCCTTGGCAAAGGGCTATGCCTTCGCCGAGGGCAAGTATGAAGCGGTGGAGTGGAAGAACAGGCCGGAAATCTCCGGCGCCGGTGCCATTATCAGCACTGTGGAAGACTACGCGAAATGGATCTACGCACTACTGAATCAGAGCGGGTTGCCTCTTTCCTCGGAGGGCTACGGAACACTCTGGACCGCGAGAGCACTCATTCCAAATTCCGAGCCGTTTTTGGCCCCGATGGCGTATGCCCTTGGATGGGACCGTTACATCTACCAGGGAGTGGAGATCATCACTCACGATGGGGGGATAGAGGGATTCGGAGCCGAGATTGTGATGATCCCCGCGCTTAAATTTGGGGTCATCACCATGGCTAACTCGACCTATTCGTCCAACTATGGGGGAACATGCCTGGCGTATGAGCTTATCGATTCGAAGCTGGGCATCGCTGCGGGGGACAGATTTGACTGGAAACAAAAGTGTGTCTTGTCATTCCTCTTGCTCTCCAACAAACTGACTAAATCGGGAGCAGATATGTCGATATCGTCGACCAAATGGACGCTTACAATGCAGACGCGGTCCAAATCTTCTACCCCGATCTCCCGTCTCCTCCTCTACCGGGGCCGACTCTCGCTATCGAAGCATACTCCGGCACCTACTGGCATGATGCCTACGGTCGGCTCGATCTCTCCATAGATGGCACCGGCACAAAGCTCCACGCGGACCGAACTAATTGTACCAATACCTGCTCTCTTACATTCGAGAACGTCACGGGAAACTACTTTATTATTAGGCTGCTTGTAGTTGGAGCCGAGACCGTCCTTCCAGCTGAGTTTTCGGTTGACCCGGACGGGAAGCCAAGGTCTGTCGGTATAGGGTGGGAGCCAACGATGGGGATAGAGAAGAAGATTTGGATGCGCAAGGTCGATGGTGATGATACCGTTGGGCTGCATCGCGCCCTGCCGTACAAGCCATCCAGAACTCCCCAGCTGCCGGAGTTTTTGACTGGCCACCTTTTTGTCTAG | MAMLVEDNDKYPGVQWTTPVNELIREDFVLDNHWTTANVTIEDILSHRTGMPGHNFSLGAVYADKQATVQDVVRSLRFLPSTAPPRTTYQYNNAMYIVASHLIQTVMGDDLGSIFQKHIWDPLGMSSTYFRLDDALASQKPLAKGYAFAEGKYEAVEWKNRPEISGAGAIISTVEDYAKWIYALLNQSGLPLSSEGYGTLWTARALIPNSEPFLAPMAYALGWDRYIYQGVEIITHDGGIEGFGAEIVMIPALKFGVITMANSTYSSNYGGTCLAYELIDSKLGIAAGDRFDWKQKYVDIVDQMDAYNADAVQIFYPDLPSPPLPGPTLAIEAYSGTYWHDAYGRLDLSIDGTGTKLHADRTNCTNTCSLTFENVTGNYFIIRLLVVGAETVLPAEFSVDPDGKPRSVGIGWEPTMGIEKKIWMRKVDGDDTVGLHRALPYKPSRTPQLPEFLTGHLFV |

A black and white chart

Description automatically generated

**Figure 1S:** A domain sequence alignment comparing UngA, UngA’, and UngA’’.

A black and white grid with many squares

Description automatically generated

**Figure 2S:** C domain sequence alignment comparing UngA, UngA’, and UngA’’.

A green text with black text

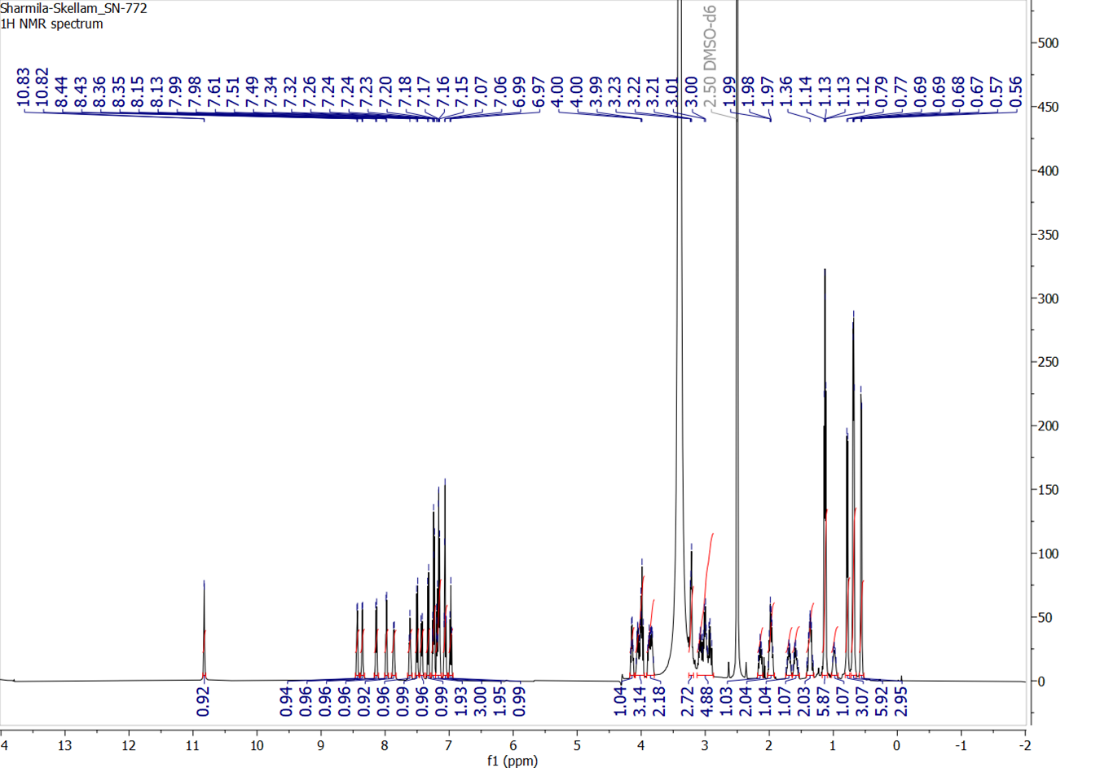
Description automatically generated

**Figure 3S:** A domain AMP binding motif identified in extracted A domains.

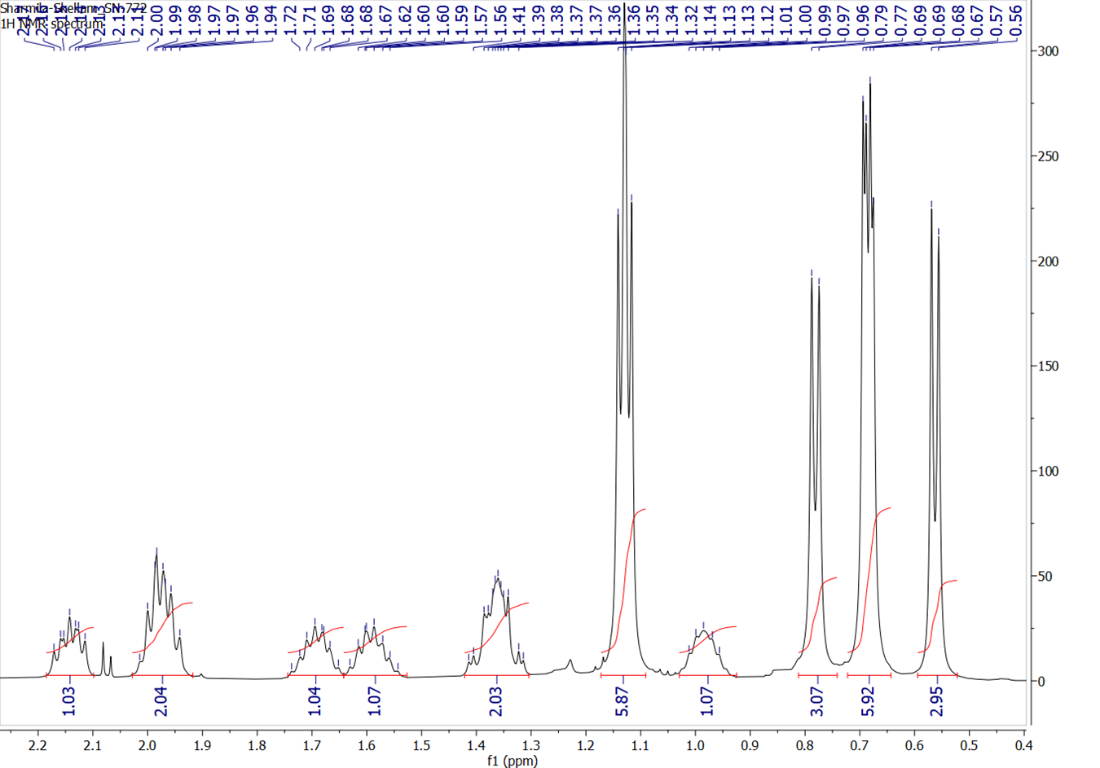
A close-up of a logo

Description automatically generated

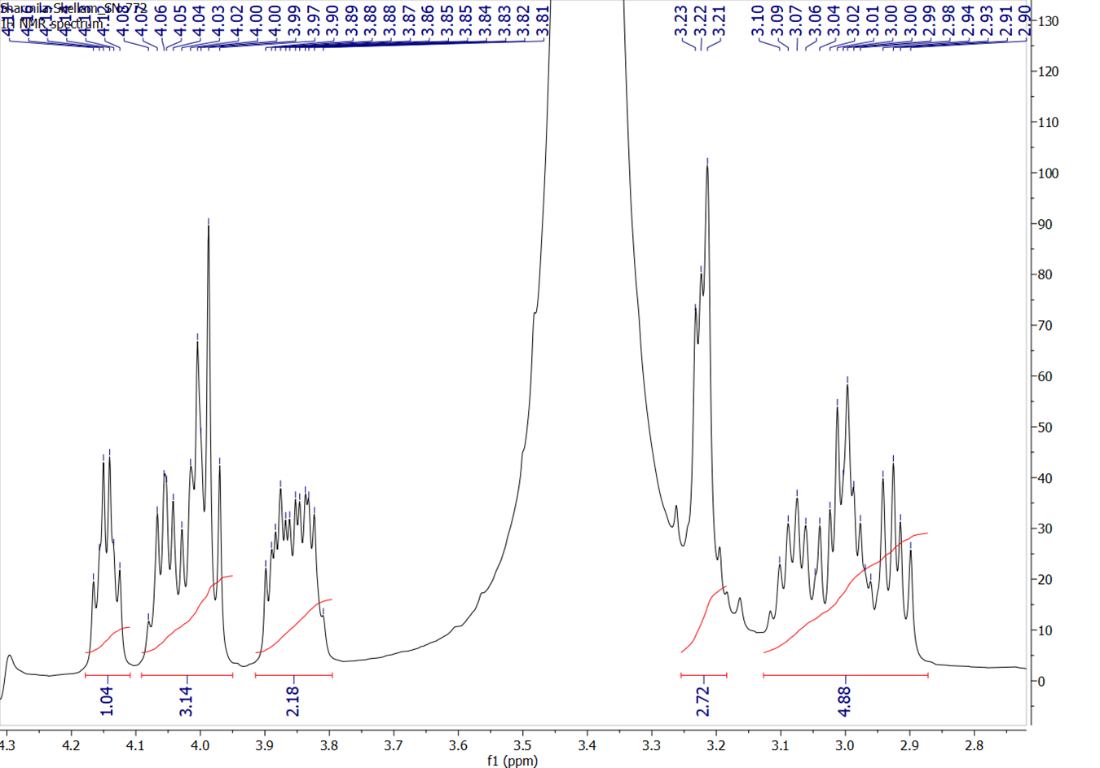
**Figure 4S:** C domain active site identified in extracted C domains*.*



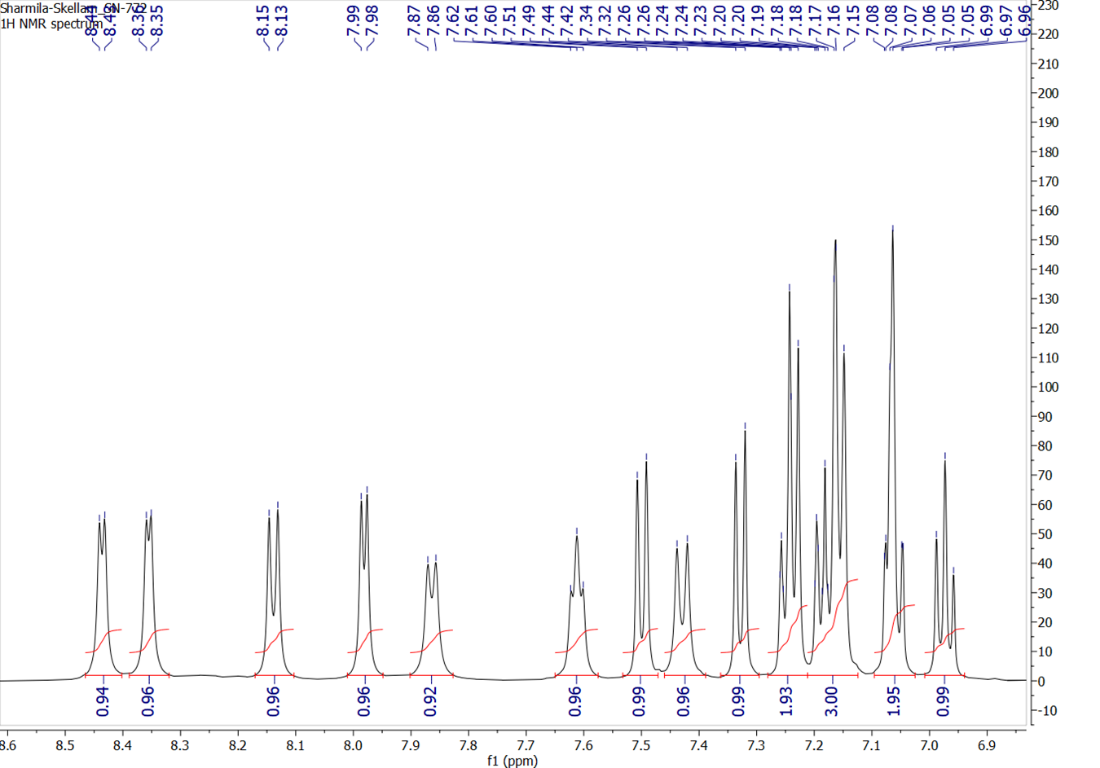
**Figure 5S:** 1H NMR spectrum of **1** (500 MHz, DMSO-*d*6).



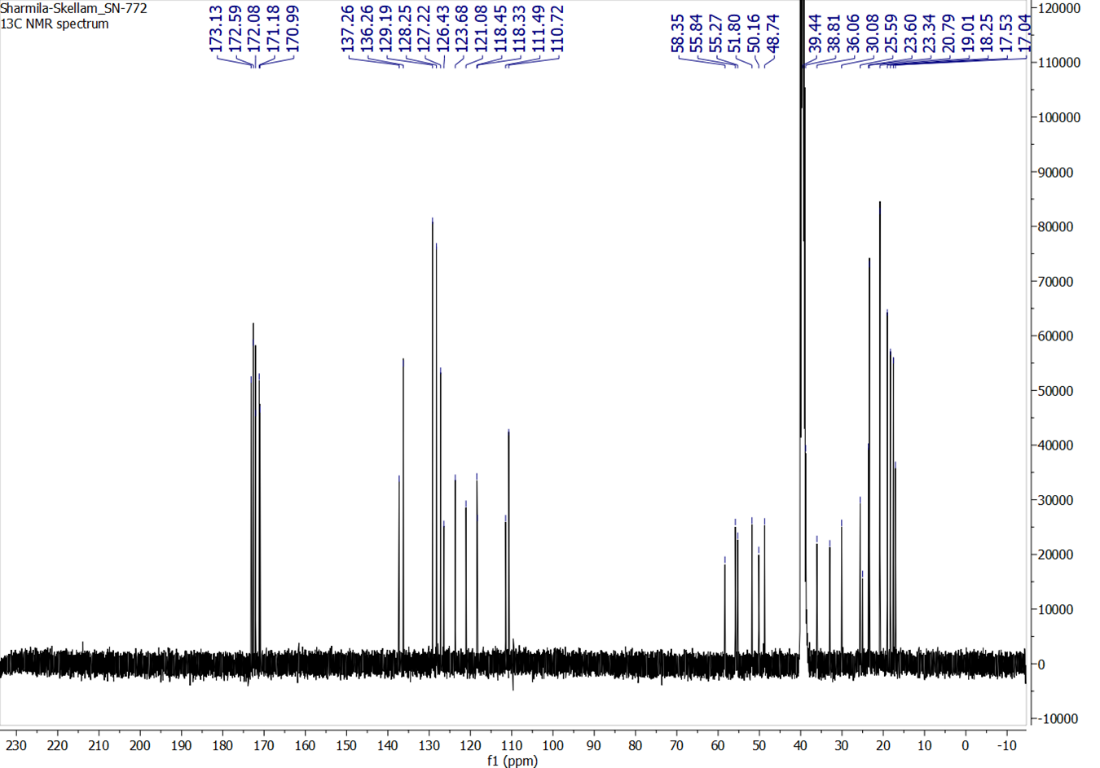
**Figure 6S:** Expansion of the 1H NMR spectrum of **1** (500 MHz, DMSO-*d*6).



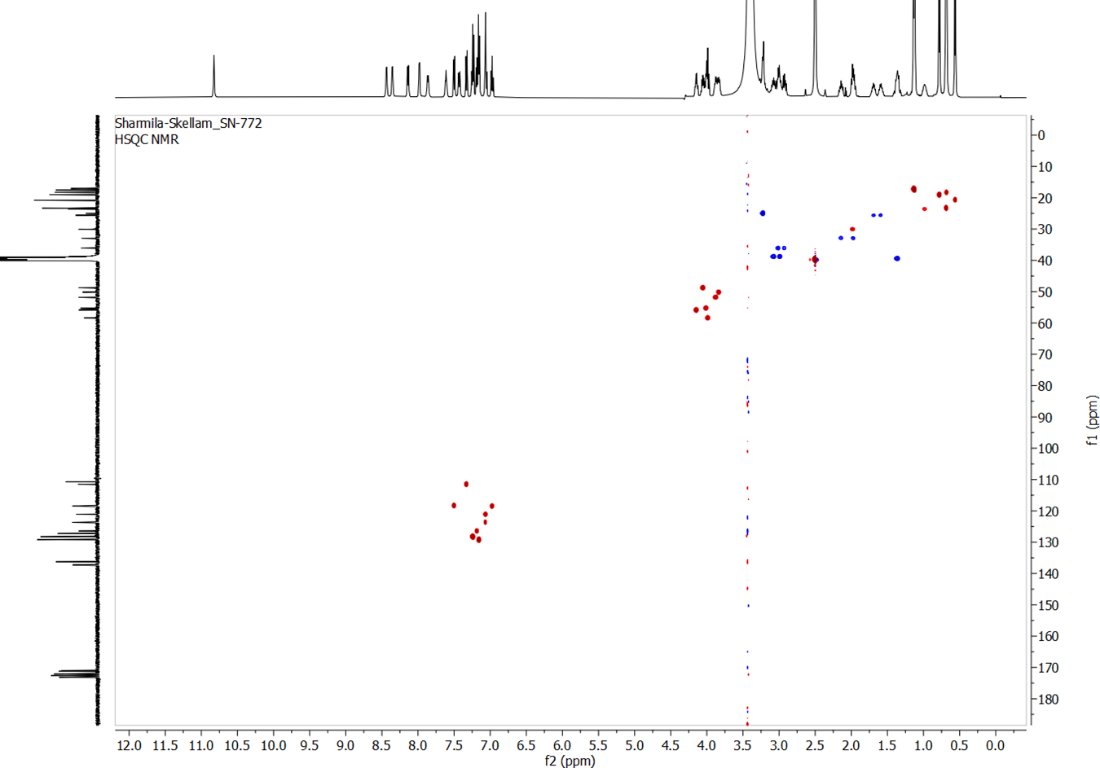
**Figure 7S:** Expansion of the 1H NMR spectrum of **1** (500 MHz, DMSO-*d*6).



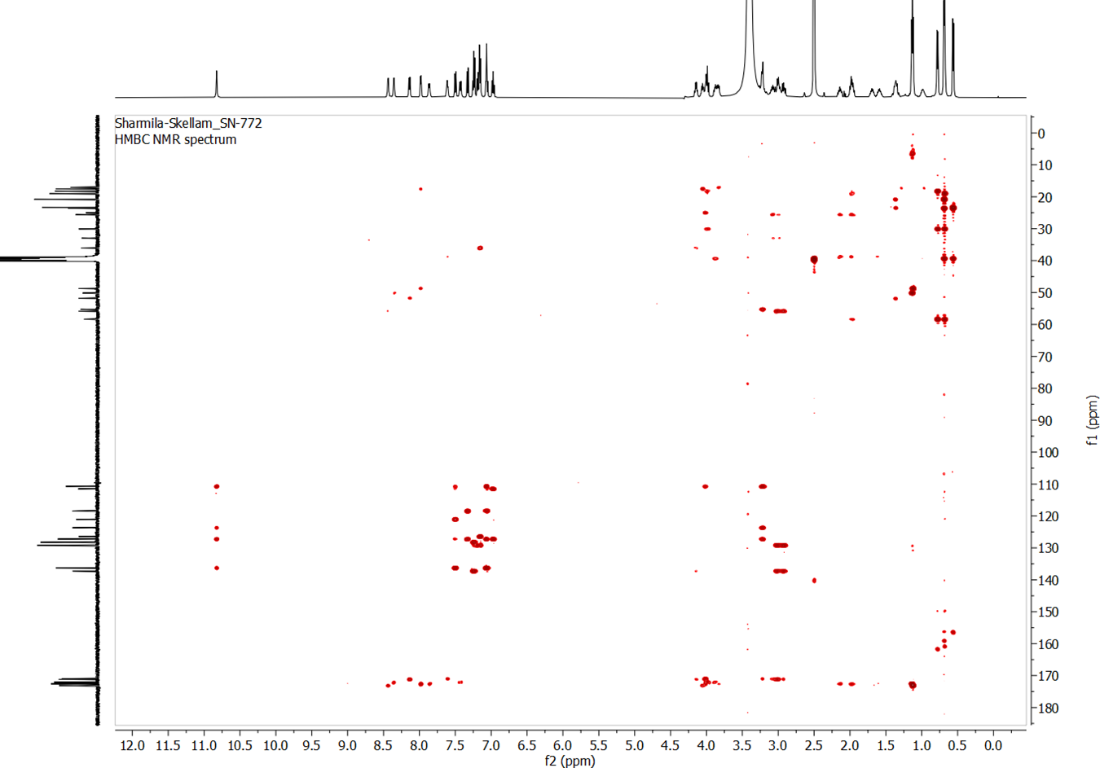
**Figure 8S:** Expansion of the 1H NMR spectrum of **1** (500 MHz, DMSO-*d*6).



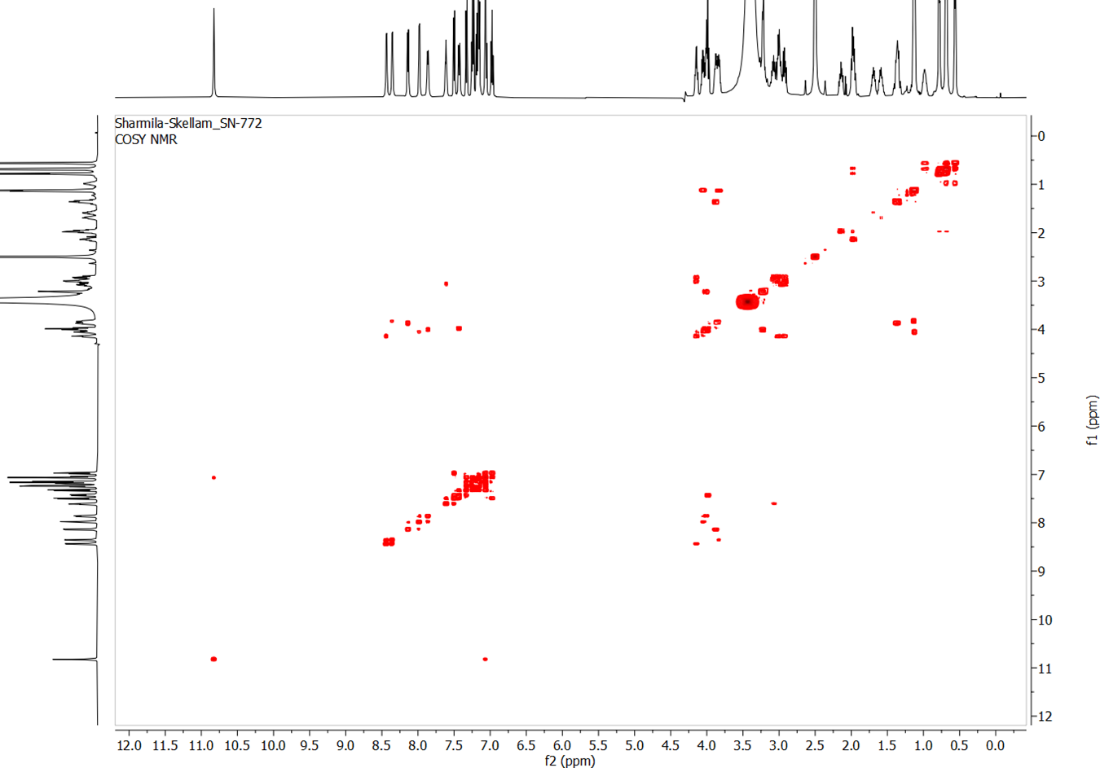
**Figure 9S:** 13C NMR spectrum of **1** (125 MHz, DMSO-*d*6).



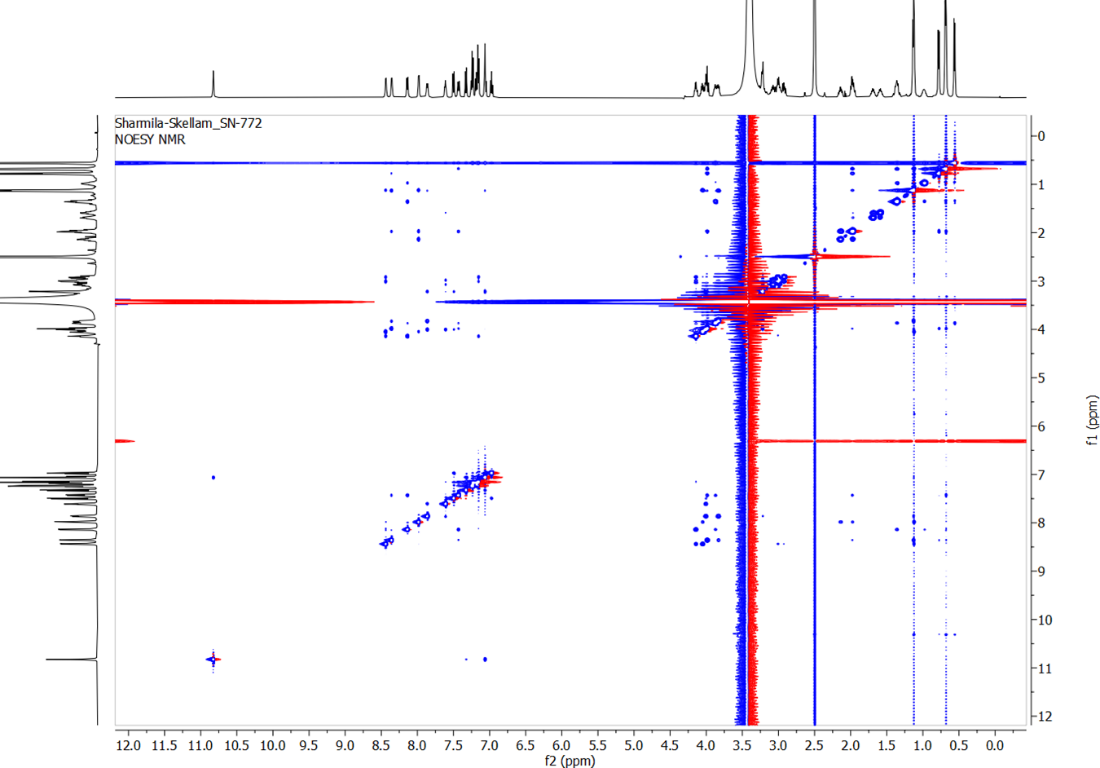
**Figure 10S:** *g*HSQC spectrum of **1** (1H:500 MHz; 13C:125 MHz, DMSO-*d*6).



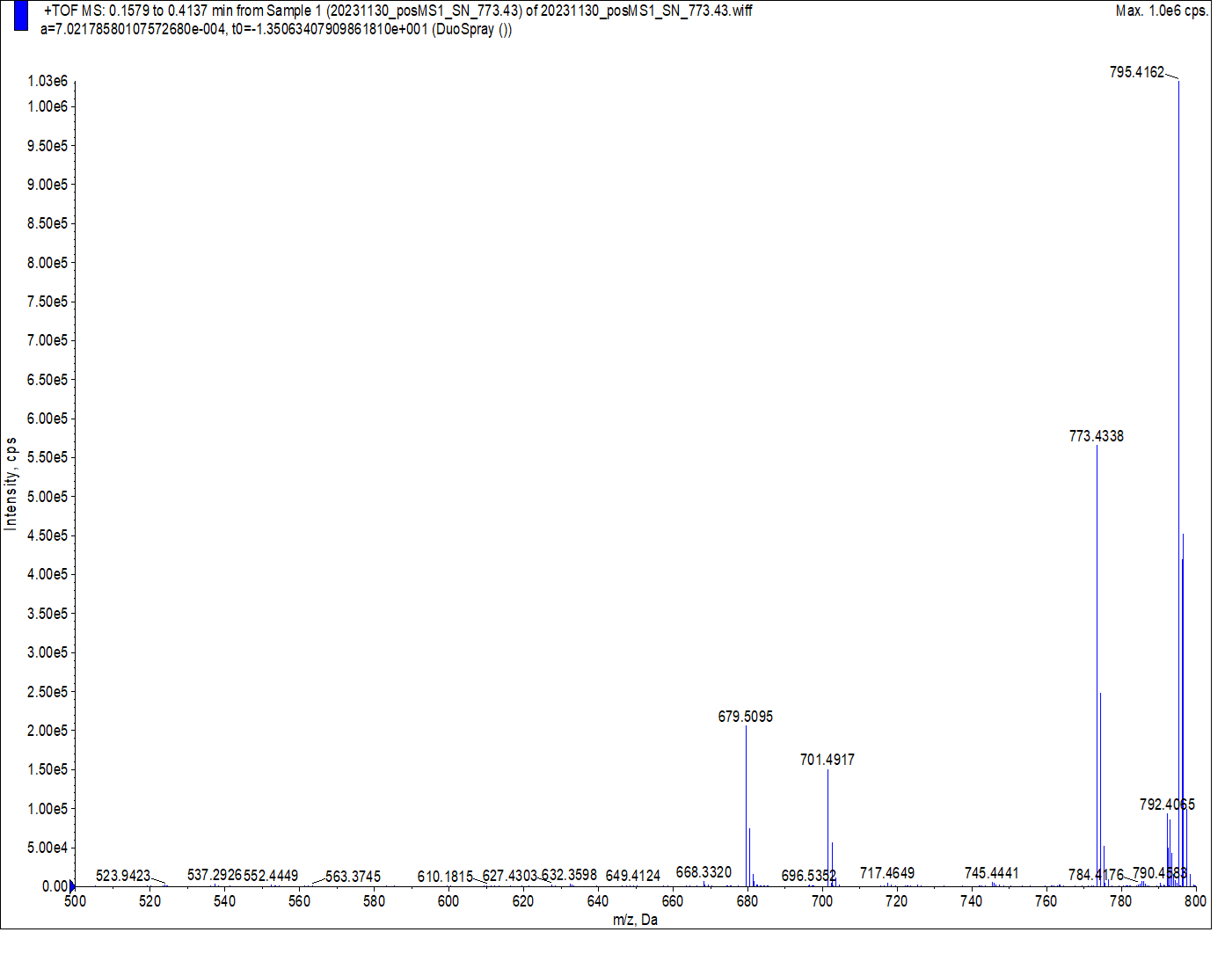
**Figure 11S:** *g*HMBC spectrum of **1** (1H:500 MHz; 13C:125 MHz, DMSO-*d*6).



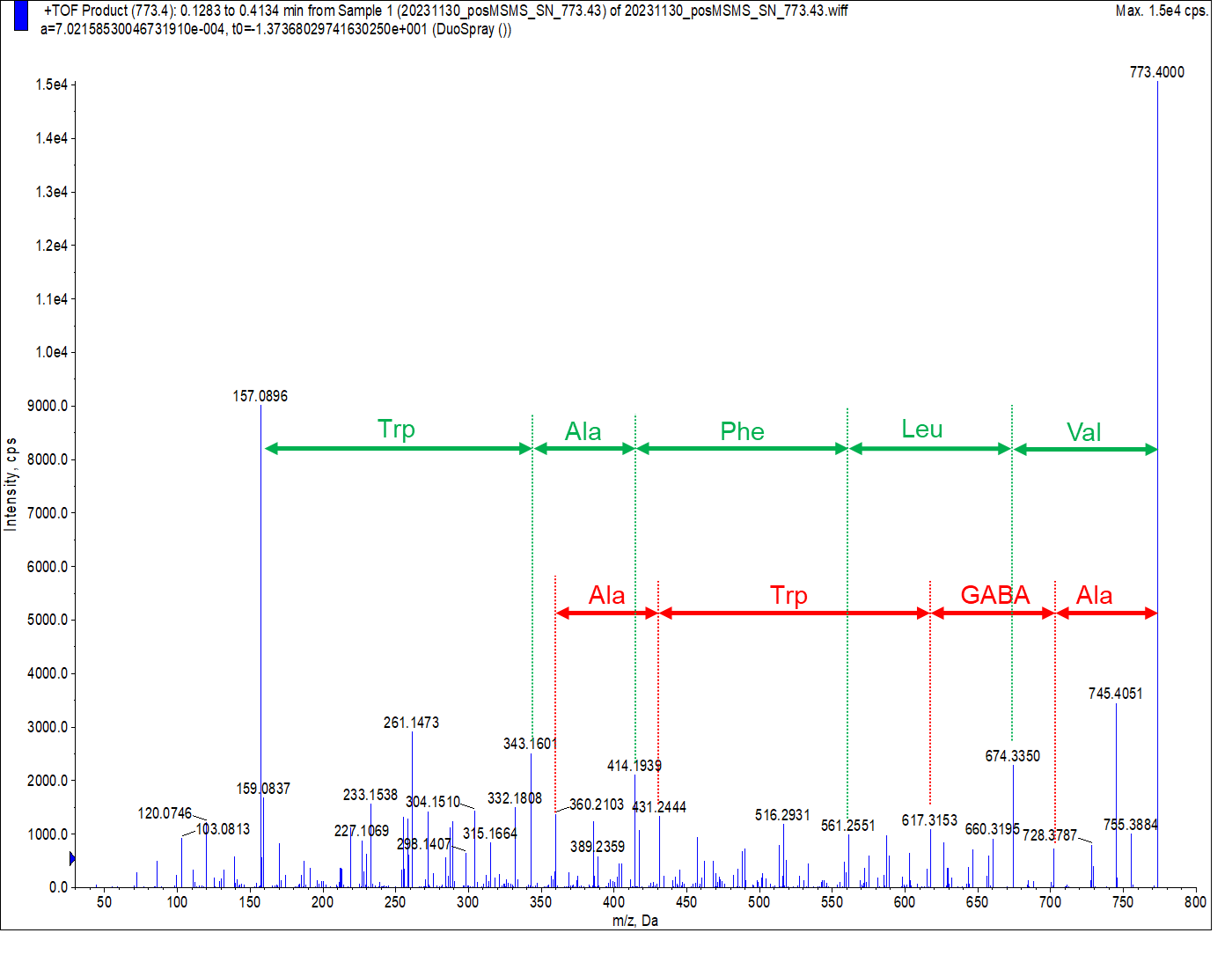
**Figure 12S:** *g*COSY spectrum of **1** (500 MHz, DMSO-*d*6).



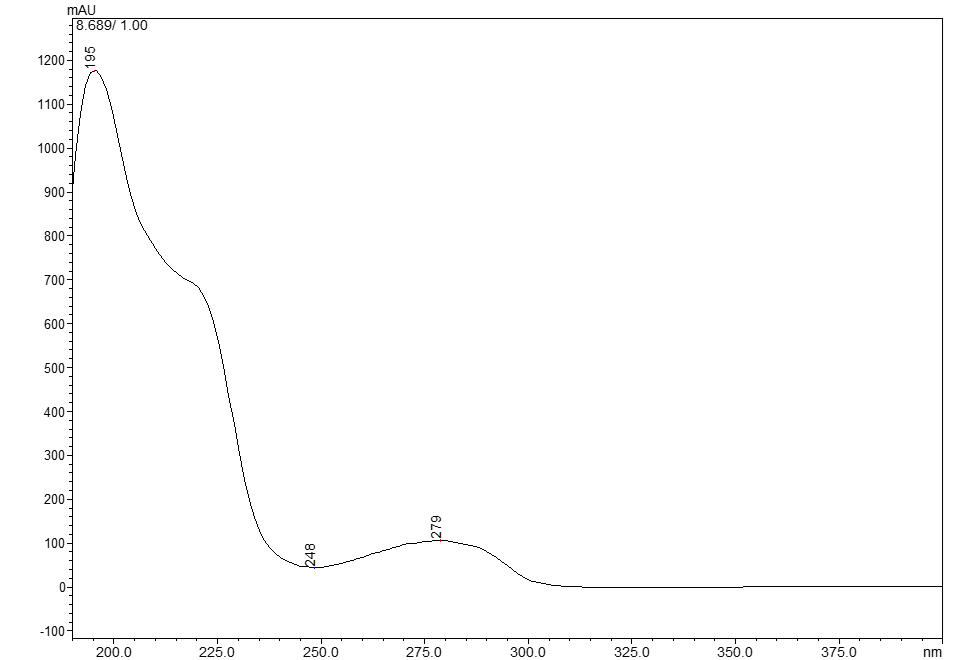
**Figure 13S:** 2D-NOESY spectrum of **1** (500 MHz, DMSO-*d*6).



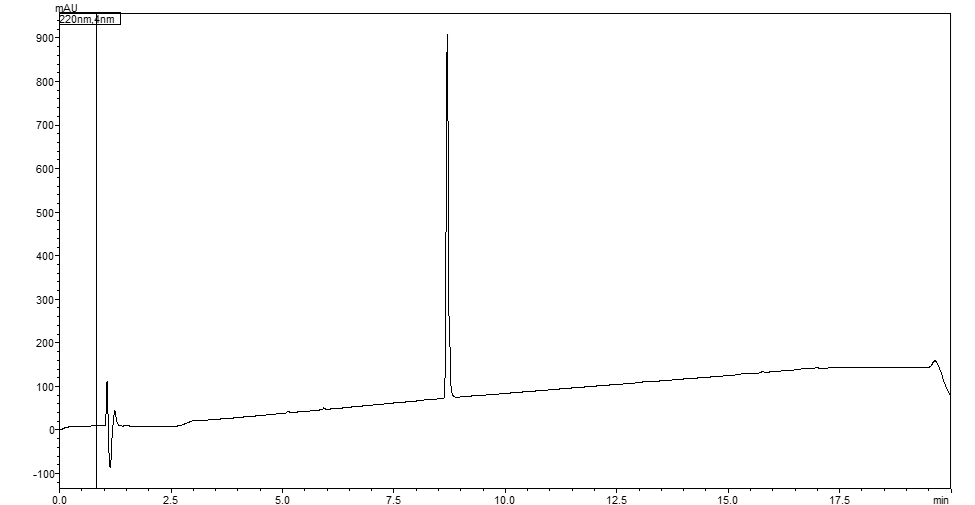
**Figure 14S:** HRESIMS spectrum of **1** in positive ionization mode.



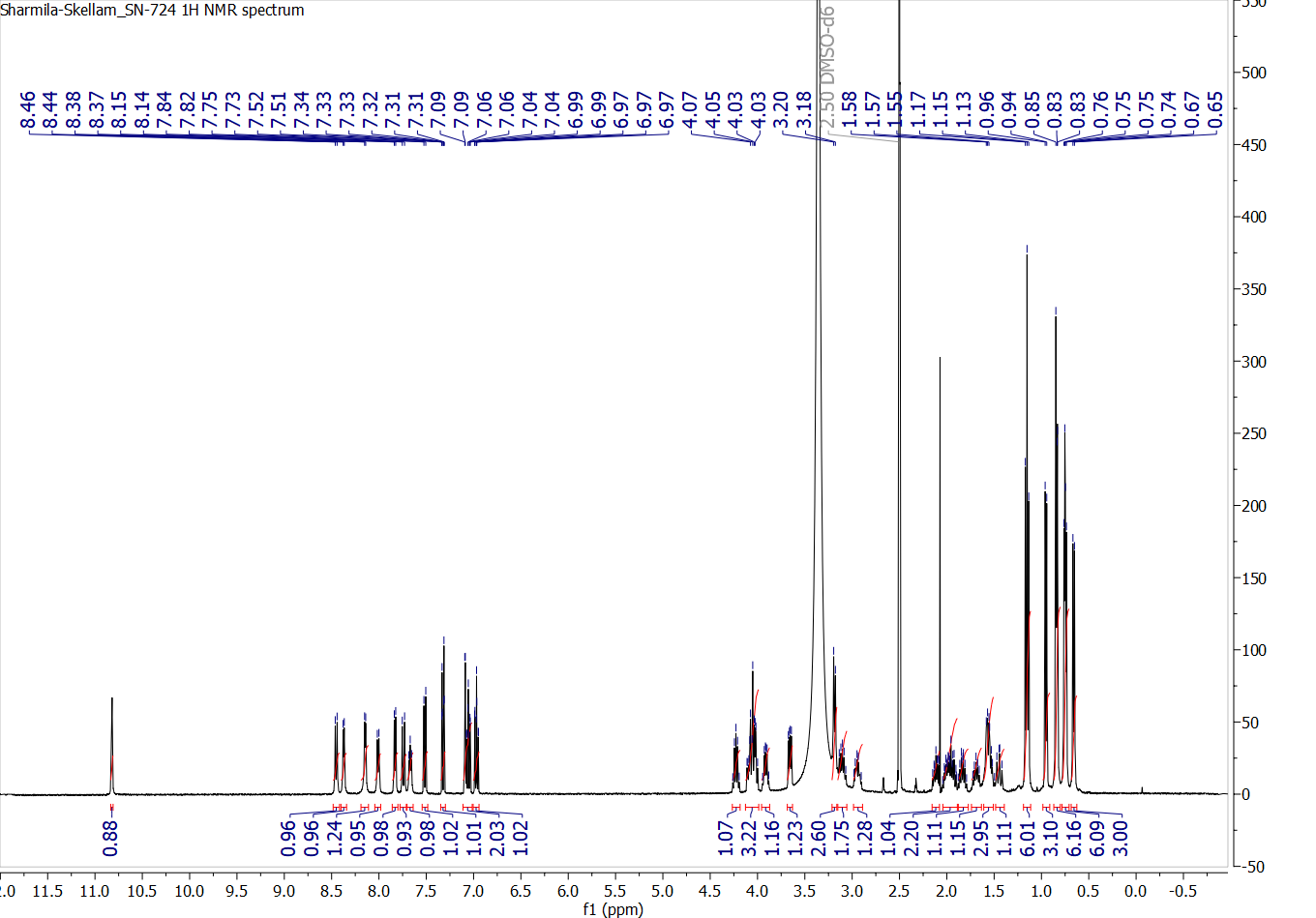
**Figure 15S:** HRESIMS/MS spectra of **1** and fragmentation patterns in positive ionization mode.



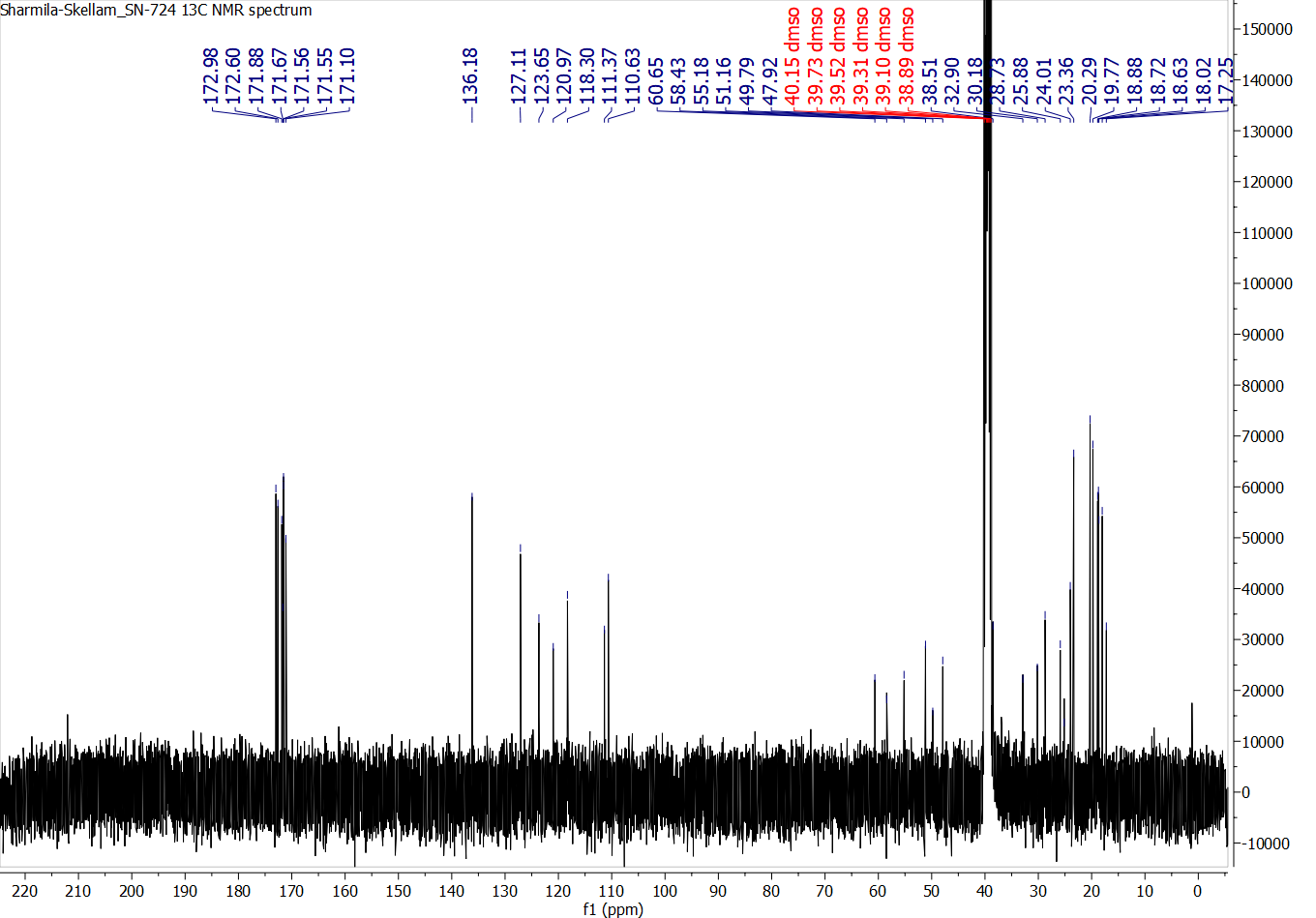
**Figure 16S:** UV spectrum (photodiode array, H2O:MeCN) of **1**.



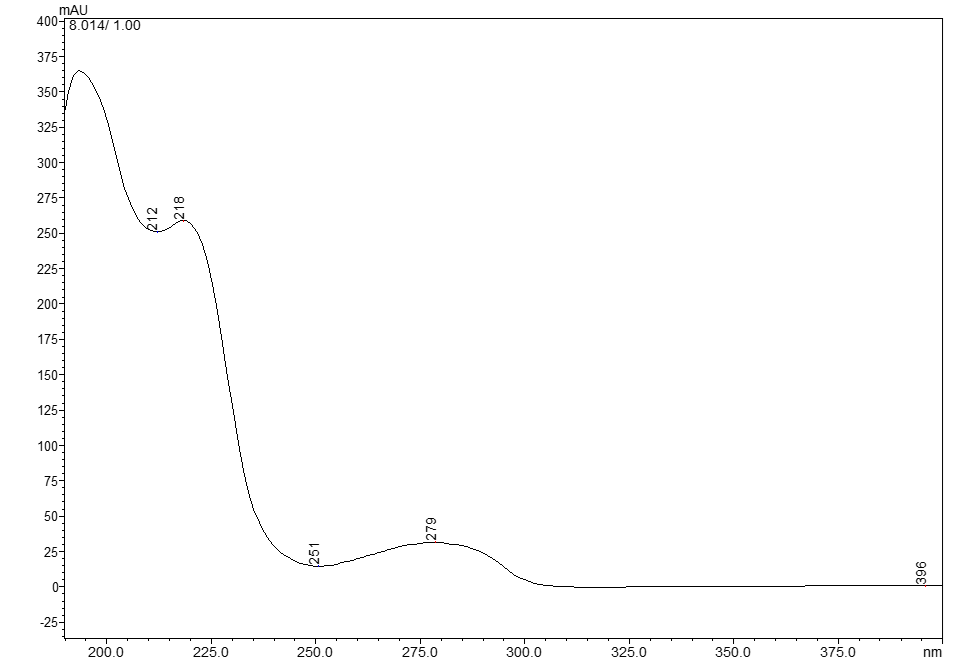
**Figure 17S:** HPLC-PDA chromatogram of unguisin J (**1**).



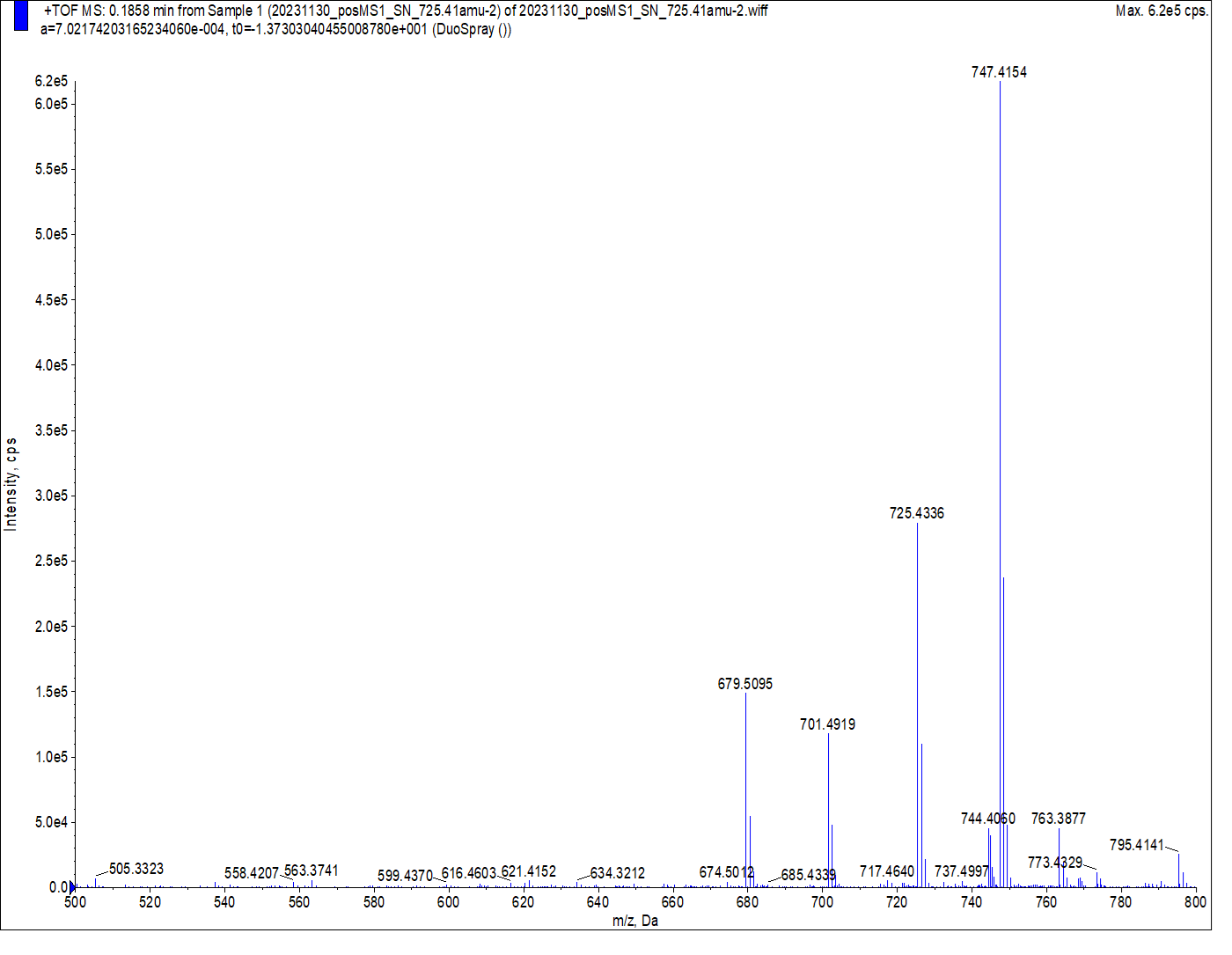
**Figure 18S:** 1H NMR spectrum of **2** (500 MHz, DMSO-*d*6).



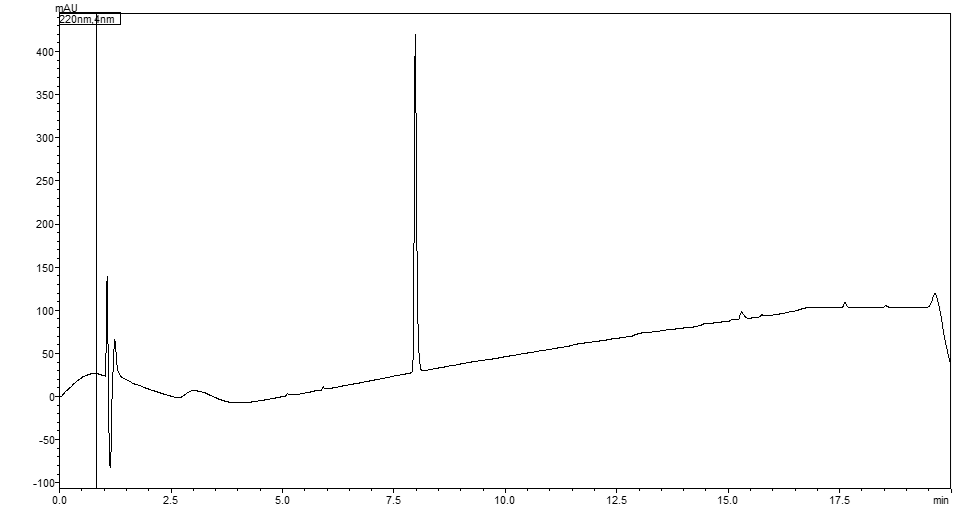
**Figure 19S:** 13C NMR spectrum of **2** (125 MHz, DMSO-*d*6).



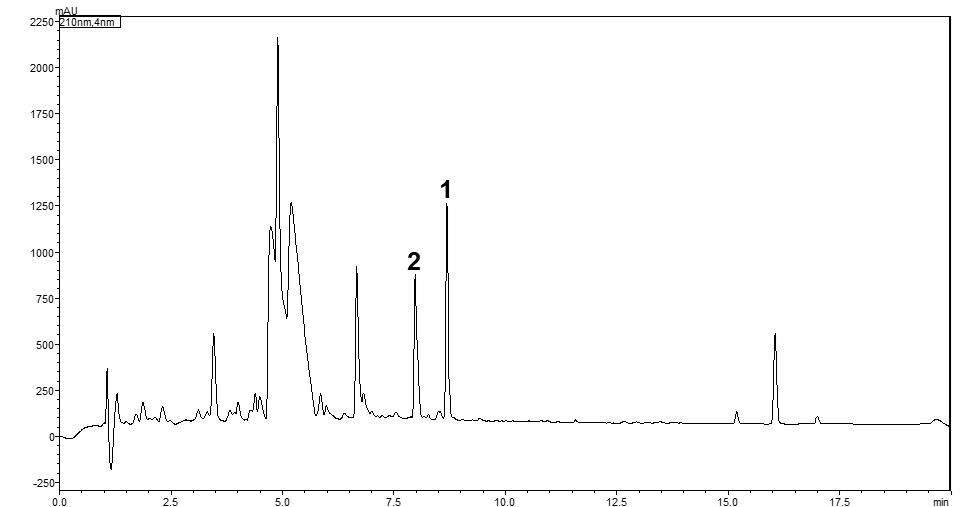
**Figure 20S:** UV spectrum (photodiode array, H2O:MeCN) of **2**.



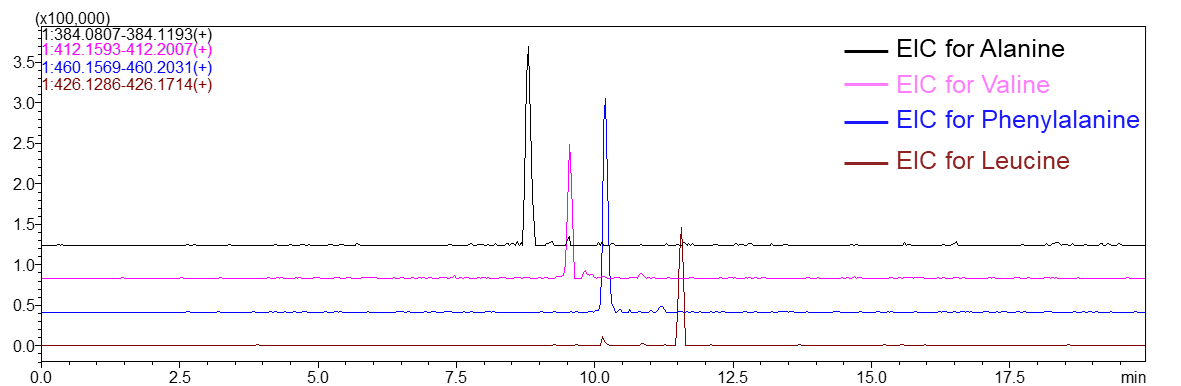
**Figure 21S:** HRESIMS spectrum of **2** in positive ionization mode.



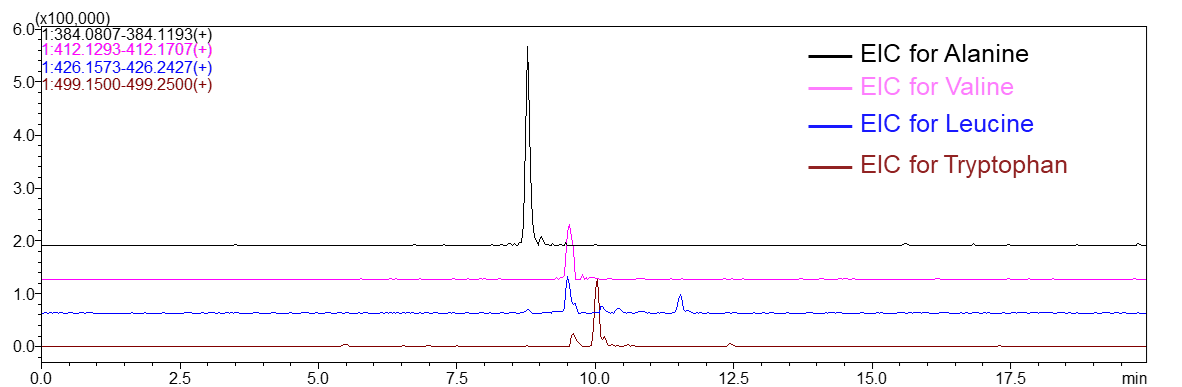
**Figure 22S:** HPLC-PDA chromatogram of unguisin B (**2**).



**Figure 23S:** HPLC-PDA chromatogram of soluble-organic extract from rice culture of *A. heteromorphus*. Peak labels represent tUnguisin J (**1**) and Unguisin B (**2**).

****

**Figure 24S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of the hydrolysate of **1**.



**Figure 25S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of the hydrolysate of **2**.



**Figure 26S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of D-alanine.



**Figure 27S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of L-alanine.



**Figure 28S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of D-valine.



**Figure 29S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of L-valine.



**Figure 30S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of D-phenylalanine.



**Figure 31S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of L-phenylalanine.



**Figure 32S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of D-leucine.



**Figure 33S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of L-leucine.



**Figure 34S:** LC-MS analysis of Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide derivative of D-tryptophan.