SUPPORTING INFORMATION

Potent human alpha-amylase inhibition by the β-defensin-like protein Helianthamide.

Christina Tysoe*, Leslie K. Williams*, Robert Keyzers, Nham T. Nguyen, Chris Tarling, Jacqueline Wicki, Ethan D. Goddard-Borger, Adeleke H. Aguda, Suzanne Perry, Leonard J. Foster, Raymond J. Andersen, Gary D. Brayer, Stephen G. Withers

a Centre for High-Throughput Biology, Michael Smith Laboratories, 185 East Mall, Vancouver, BC V6T 1Z4 Canada

b Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada

c Department of Chemistry, University of British Columbia, 2036 Main Mall, Vancouver, British Columbia V6T 1Z1, Canada

d Department of Earth and Ocean Sciences, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada

Table of Contents

1. Materials and Methods S2
2. MS/MS data of tryptic digests of natural helianthamide (Fig. S1 and S2) S4
3. MS/MS data of selected peaks for disulfide analysis of helianthamide (Fig. S3) S6
4. N-terminal region of helianthamide (Fig. S4) S7
5. Disulfide bridges in the helianthamide/PPA crystal structure (Fig. S5) S8
6. Sequence of the barnase'-helianthamide gene (Fig. S6) S9
7. Peptide sequence of the barnase'-helianthamide fusion protein (Fig. S7) S10
8. Secondary structural analysis of CD spectra (Table S1) S11
9. Table of LC-MS/MS peaks and corresponding sequences (Table S2) S12
10. Specificity profile of helianthamide (Table S3) S13
11. X-ray crystallography structure determination statistics (Table S4) S14

* These two authors contributed equally

1 To whom correspondence should be addressed
MATERIALS AND METHODS

Collection and extraction of animal material. *Stichodactyla helianthus* was collected by hand at a depth of 2-3 meters at Cape Capucin, Island of Dominica. Specimens were immediately frozen and transported to Vancouver, British Columbia. Prior to analysis, the *S. helianthus* were exhaustively extracted with MeOH (3 x 300 mL).

Activity guided isolation. The *S. helianthus* methanolic extract from a 154 g *S. helianthus* specimen was further separated on a Diaion HP-20 column, which was eluted with 150 mL portions of 0%, 30%, 50%, 75%, and 100% acetone in water. Greatest activity was traced to the 30% acetone/water fraction (162 mg), which was then chromatographed on a Sephadex LH-20 size exclusion column (28 x 710 mm) eluting with 25% methanol/75% water/0.1% trifluoroacetic acid (TFA). The material (13.1 mg) was then purified using RP-HPLC with a C8 column eluting with 23% acetonitrile/77% water/0.1% TFA to yield the active peptide (2.0 mg).

Peptide sequencing. Disulfide linkages were reduced with dithiothreitol (100 mM) for 1 hour at 56°C after which the resulting free cysteines were alkylated with iodoacetamide (100 mM) for 1 hour at 25°C in pH 7.5 ammonium bicarbonate buffer (50 mM). MALDI-TOF analysis indicated a change in mass of 350.4 after reduction and alkylation which corresponds to approximately six alkylated cysteines. It could be deduced that the native material possessed three disulfide bonds. Peptide digestions were carried out by incubating reduced and alkylated peptide with trypsin (Promega Gold) in pH 8.5 Tris buffer (50 mM) for 18-24 hours at 37 °C, or with GluC (V8 protease, Roche) in pH 8.5 sodium phosphate buffer (50 mM) for 18-24 hours.

De novo sequencing: Sequencing was performed by LC-MS/MS on an LC Packings capillary LC system (Dionex) coupled to a QSTAR Pulsar quadrupole TOF mass spectrometer (Applied Biosystems). MS data were acquired using Analyst QS software (Applied Biosystems) for information-dependent acquisition. N2 was used as collision gas and the ionization tip voltage was 2200 V.

Edman sequencing: Digested peptides were separated on a Vydac HPLC microbore C_{18} column (ODS-300, 7 µm, 1 x 50 mm). The peptide fractions were loaded on a precycled glass fiber filter coated with Biobrene and sequenced by the Edman degradation liquid phase method using a Procise cLC-494 (Applied Biosystems) equipped with an on-line 140D Phenylthio-hydantoin analyser (Applied Biosystem). The phenylthiohydantoin amino acid (PTH-aa) derivatives were determined by comparison with standards analyzed at the start of the sequence analysis.

CD spectroscopy. Analysis was conducted on a Jasco J-815 CD spectrophotometer. Samples of synthetic, natural, and recombinant helianthamide were diluted in distilled water. Analysis was run from 260 to 190 nm for the natural and synthetic material, and 250 to 190 nm for the recombinant material. The resulting data was analyzed on the Dichroweb server(1). All three data sets were analyzed with CONTIN software using reference sets 4, 7, and SP175(2,3). The data of the natural and synthetic material was also analyzed using CDSSTR software with reference sets 4, 7, and SP175. The secondary structural estimates were averaged and presented in Table S1.
**Thiol titration assay.** A solution of thiol titration buffer was prepared (0.2 mM 5,5'-dithiobis-(2-nitrobenzoic acid), 6 M guanidinium chloride, 20 mM HEPES, 1 mM EDTA, pH 7.4). Helianthamide was added to the solution in final protein concentrations ranging from 2 – 6 µM(4). The release of 2-nitro-5-thiobenzoate was monitored at 412 nm at 25 ºC over the course of 30 minutes. During each measurement, ΔA_{412} was also measured for the thiol titration buffer to take into account the rate of spontaneous hydrolysis of the DTNB reagent. The experiment was also conducted with BSA, which has one free thiol group, as a positive control.

**Disulfide assignment of recombinant material.** Intact helianthamide was incubated in 11 M hydrochloric acid at 37 ºC for 4 days. The sample was diluted and neutralized before desalting by C18 reverse phase chromatography. The peptide mixture was injected into an Agilent 6460 QQQ LC/MS mass spectrometer. Elution and MS/MS analysis was carried out over a 40 minute period. A script was written to predict all possible fragments for the helianthamide sequence in the event of random hydrolysis and was used to analyze the resulting MS and MS/MS peaks.

**Expression of human pancreatic α-amylase.** A detailed protocol for the expression and purification of HPA can be found in the work done by Rydberg et al. (5) Expression was carried out in *Pichia pastoris*. Colonies were grown in 60 mL BMGY media at 30ºC 200 RPM. 20 mL of the overnight culture was added to 600 mL BMGY. After 16 hours the cells were transferred to 300 mL BMMY. After one day 2 mL of 50% methanol was added. The culture was left overnight. The culture supernatant was collected and purified via Phenyl Sepharose and Hitrap Q columns. Deglycosylation was performed with *N*-glycosidase F.

1. Whitmore, L.; Wallace, B. A. Protein secondary structure analyses from Circular Dichroism spectroscopy: methods and reference databases. *Biopolymers* 2008, 89(5), 392-400.
2. Sreerama, N.; Woody, R. W. Estimation of protein secondary structure from CD spectra: comparison of CONTIN, SELCON and CDSSTR methods with an expanded reference set. *Anal Biochem* 2000, 287, 252-260.
3. Lees, J. G.; Miles, A. J.; Wien, F.; Wallace, B. A. A reference database for circular dichroism spectroscopy covering fold and secondary structure space. *Bioinformatics* 2006, 22, 1955-1962.
4. Riener, C.; Kada, G.; Gruber, H. Quick measurement of protein sulfhydryls with Ellman’s reagent and with 4,4’-dithiodipyridine. *Anal Bioanal Chem* 2002, 373, 266-276.
5. Rydberg, E.H.; Sidhu, G.; Vo, H. C.; Hewitt, J.; Cote, H. C.; Wang, Y.; Numao, S.; MacGillivray, R. T.; Overall, C. M.; Brayer, G. D.; Withers, S. G. Cloning, mutagenesis, and structural analysis of human pancreatic alpha-amylase expressed in Pichia pastoris. *Protein Sci* 1999, 8(3), 635-643.
Figure S1. MS/MS data of tryptic digests of natural helianthamide.
Figure S2. MS/MS data of tryptic digests of natural helianthamide (continued).
**Figure S3.** MS/MS data of selected peaks and their corresponding sequences for disulfide analysis of helianthamide. The data confirm the 1-5,2-4,3-6 connectivity seen in the crystal structure.
Figure S4. N-terminal region of helianthamide.
Figure S5. Omit maps of disulfide bridging in the helianthamide/PPA crystal structure.
ATG AAA CAA TCC ACC ATC GCA CTG GCC CTG CTG CCG CTG CTG
TTC ACG CCG GTT ACA AAG CCG CTG GTG CAT CAT CAT CAT CAT
CAC TCG AGT GGC GCA CAG GTT ATT AAC ACC TTT GAT GGT GTT
GCT GAC TAT CTG CAA ACG TAC CAT AAA CTG CCG GAT AAT TAT
ATC ACC AAA TCA GAA GCA CAG GCC CTG GGT TGG GTC GCA TCG
AAA GGT AAC CTG GCA GAT GTG GCT CCG GGC AAA AGT ATT GGC
GCT GAC ATC TCC TCC AAT CGT GAA GGT AAA CTG CCG GGC AA
TCT GGT CGT ACC TGG CGC GAA GCG GAT ATT AAC TAT ACG TCA
GGC TTT CGT AAT TCG GAT CGC ATT CTG TAC AGC TCT GAC TGG
CTG ATC TAT AAA ACC ACG GAC GCC TAC CAA ACC TTC ACG AAA
ATT CGT ATC GAA GGC CGC GAA AGT GGT AAC TCC TGC TAT ATT
TAC CAC GCC GTT AGC GGT ATC TGC AAA GCG TCT TGT GCC GAA
GAT GAA AAA GCA ATG GCA GGC ATG GCC GTG TGT GAA GGT CAT
CTG TGT TGT TAC AAA ACC CCG TGG TGA TAG

Figure S6. Sequence of the barnase'-helianthamide gene optimized for expression in *E. coli.*
Figure S7. Peptide sequence of the complete barnase’-helianthamide fusion protein. His-tag is highlighted in green. The active site mutation of barnase is highlighted in pink. The TEV protease recognition site is highlighted in yellow.
Table S1. Secondary structural analysis of the CD spectra (Fig 1b) of the natural, synthetic, and recombinant helianthamide

|                | Proportion of each secondary structural estimate |
|----------------|-----------------------------------------------|
|                | α-helix | β-sheet | turns | other |
| Synthetic      | 0.06    | 0.32    | 0.18  | 0.44  |
| Natural        | 0.19    | 0.32    | 0.18  | 0.31  |
| Recombinant    | 0.11    | 0.33    | 0.23  | 0.33  |
| M     | Parent peak | Sequence               |
|-------|-------------|------------------------|
| 1198.42 | 600.21 [M+2H]^{2+} | IYHVSGI\text{C} CE   |
| 1279.58 | 640.79 [M+2H]^{2+} | HGVSGICKA MGVC     |
| 1286.70 | 429.90 [M+3H]^{3+} | GICK KAMAGMGV C     |
| 1311.70 | 656.85 [M+2H]^{2+} | CYIYHVG LCC C       |
| 1327.68 | 443.56 [M+3H]^{3+} | VC ICKAS CAE CY     |
| 1347.66 | 449.22 [M+3H]^{3+} | NSCYI LCCYK C       |
| 1416.68 | 709.34 [M+2H]^{2+} | GIC EKAMAGMGVCE     |
| 1439.60 | 480.88 [M+3H]^{3+} | SC EGHLCC Y SCAE    |
| 1534.56 | 768.28 [M+2H]^{2+} | ESGNSC GHLCC SCAE   |
| 1551.70 | 776.85 [M+2H]^{2+} | GNSC GHLCCY KASCA   |
| 1606.82 | 804.41 [M+2H]^{2+} | CY GVEGHL C SGICK   |
| 1631.78 | 816.89 [M+2H]^{2+} | SESGNSC LCCYK CAE   |
| 1698.81 | 566.27 [M+3H]^{3+} | SGNSC LCC CAEDEKAM  |
| 1760.80 | 881.40 [M+2H]^{2+} | GNSC LCCYKT SCAEDE  |
| 1818.97 | 607.32 [M+3H]^{3+} | SESGNSECYIVHG CC SCA |
Table S3. Inhibitory specificity of helianthamide

| Enzyme                                      | $K_i$     |
|---------------------------------------------|-----------|
| Porcine pancreatic $\alpha$-amylase         | 0.1 nM    |
| Human maltase-glucoamylase                  | N.I.      |
| *Roseburia inulinivorans* $\alpha$-amylase A | N.I.      |
| *Butyrivibrio fibrisolvens* $\alpha$-amylase B | N.I.      |
| *Bacillus licheniformis* $\alpha$-amylase   | N.I.      |
| Bovine liver $\beta$-galactosidase          | N.I.      |
| Green coffee bean $\alpha$-galactosidase    | N.I.      |
| Jack bean $\alpha$-mannosidase              | N.I.      |
| Brewer’s yeast $\alpha$-glucosidase         | N.I.      |
| Almond $\beta$-glucosidase                  | N.I.      |

N.I. indicates no inhibition observed at a concentration of 1 $\mu$M.
| Table S4. Summary of Structure Determination Statistics |
|-------------------------------------------------------|
| **Data Collection Parameters**                        |
| **helianthamide/ PPA Inhibitor Complex**               |
| Space group                                           |
| P2₁2₁2₁                                               |
| Unit cell dimensions (Å)                              |
| a: 43.75                                              |
| b: 103.28                                             |
| c: 111.90                                             |
| Total no. reflections collected                       |
| 116266 (8116)                                         |
| No. of unique reflections                             |
| 16310 (1177)                                          |
| Mean I/σI<sup>a</sup>                                 |
| 17.3 (7.5)                                            |
| Multiplicity<sup>a</sup>                              |
| 7.1 (6.8)                                             |
| Merging R-factor (%)<sup>a</sup>                      |
| 10.4 (29.1)                                           |
| Maximum resolution (Å)                               |
| 2.60                                                  |

| **Structure Refinement Values**                        |
| Number of reflections                                  |
| 15472                                                 |
| Resolution range (Å)                                  |
| 49.19 - 2.6                                           |
| Completeness (%)<sup>a</sup>                          |
| 95.2 (94.7)                                           |
| No. protein atoms                                      |
| 3904                                                  |
| No. inhibitor atoms                                    |
| 322                                                   |
| No. solvent atoms                                      |
| 219                                                   |
| Average thermal factors (Å<sup>2</sup>)                |
| Protein atoms                                         |
| 16.9                                                  |
| Inhibitor atoms                                       |
| 21.3                                                  |
| Solvent atoms                                         |
| 35.5                                                  |
| Overall                                               |
| 17.2                                                  |
| Final R-free value (%)<sup>b</sup>                    |
| 22.6                                                  |
| Final R-factor (%)                                     |
| 18.1                                                  |

| **Structure Stereochemistry**                          |
| r.m.s. deviations                                      |
| bonds (Å)                                             |
| 0.004                                                 |
| angles (°)                                            |
| 0.882                                                 |

<sup>a</sup> Values in parentheses refer to the highest resolution shell: 2.73-2.6 Å.  
<sup>b</sup> 5% of the data was set aside to calculate R-free.