Wade Research Foundation Reports (2010) 5 (3)

Searching for Amelia Earhart at the Molecular Level: Peptide AMELIAEARHART

David Wade
Wade Research Foundation, P.O. Box 257, Princeton, New Jersey 08542-0257, USA

Figure 1. Amelia Earhart in Newark, N.J. (Photo courtesy of the Atchison Globe.)

Abstract

This article proposes the synthesis and study of a peptide that is composed of 13 amino acids, and that has an amino acid composition and sequence that can be described with International Union of Pure and Applied Chemistry-International Union of Biochemistry and Molecular Biology, Joint Commission on Biochemical Nomenclature single letter abbreviations corresponding to the sequence of letters in the name of Amelia Earhart, the famous aviatrix of the early 20th century: Alanine (A)-Methionine (M)-Glutamic acid (E)-Leucine (L)-Isoleucine (I)-Alanine (A)-Glutamic acid (E)-Alanine (A)-Arginine (R)-Histidine (H)-Alanine (A)-Arginine (R)-Threonine (T). The existence of similar amino acid sequences in nature, and possible structures and uses for the peptide are discussed. The peptide would probably exhibit biological activities.

Introduction

Amelia Earhart

Amelia Earhart (1897-1937) was among the most famous women aviators of the 20th century [1, 2]. She was born on July 24, 1897, and raised in Atchison, Kansas. Among her many pioneering achievements, Earhart was the first woman to fly across the Atlantic Ocean (1928), to fly solo across the Atlantic (1932), to fly across the Pacific Ocean (1935), and to receive the Distinguished Flying Cross. In Figure 1, Earhart is shown leaning on the wheel housing of an airplane at Newark, New Jersey, and waving to the photographer. The photo was taken after Earhart had completed a 19 hour, solo, nonstop, transcontinental flight from Los Angeles, CA, to Newark on August 25, 1932. She was the first woman to make such a flight. Earhart disappeared somewhere over the Pacific Ocean during a round-the-world flight on July 2, 1937.

This article proposes the synthesis and study of a peptide with an amino acid (AA) sequence that can be represented by the International Union of Pure and Applied Chemistry-International Union of Biochemistry and Molecular Biology, Joint Commission on Biochemical Nomenclature (IUPAC-IUB, JCBN) single letter abbreviations for AAs [Table 1] that correspond to the letter sequence comprising the name, Amelia Earhart: Alanine (A)-Methionine (M)-Glutamic acid (E)-Leucine (L)-Isoleucine (I)-Alanine (A)-Glutamic acid (E)-Alanine (A)-Arginine (R)-Histidine (H)-Alanine (A)-Arginine (R)-Threonine (T). The hyphens between consecutive AAs in the sequence represent covalent chemical bonds, called amide or peptide bonds (i.e., -CO-NH-).

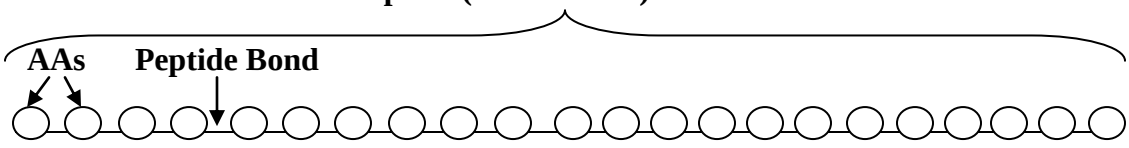
AAs, peptides and proteins

Peptides are polymers of AAs that are often compared to beads on a string, where the beads are AAs and the string is covalent chemical bonds, called peptide or amide bonds, that link successive AAs (Figure 2) [4]. AA polymers containing less than 100 AAs are peptides, and those containing 100 or

Table 1. The IUPAC-IUB, JCBN one- and three-letter abbreviations for amino acids (AAs) [3]. The letters B, J, O, X, and Z, marked with asterisks, are either unassigned, or have ambiguous assignments.

1-letter symbol	3-letter symbol	AAs	1-letter symbol	3-letter symbol	AAs
A	Ala	Alanine	N	Asn	Asparagine
B*	Asx	Aspartic acid or Asparagine	**O***	(None)	(None)
C	Cys	Cysteine	P	Pro	Proline
D	Asp	Aspartic acid	Q	Gln	Glutamine
E	Glu	Glutamic acid	R	Arg	Arginine
F	Phe	Phenylalanine	S	Ser	Serine
G	Gly	Glycine	T	Thr	Threonine
H	His	Histidine	U	Sec	Selenocysteine
I	Ile	Isoleucine	V	Val	Valine
J*	(None)	(None)	W	Trp	Tryptophan
K	Lys	Lysine	**X***	Xaa	Unknown or 'other' amino acid
L	Leu	Leucine	Y	Tyr	Tyrosine
M	Met	Methionine	**Z***	Glx	Glutamic acid or Glutamine

Figure 2. The relationship between AAs, peptides, and proteins [4].



(Continued from page 1.)

more AAs are proteins. Peptides are ubiquitous in nature, where they perform functions essential for life. An example of a well known peptide is the hormone, insulin, a polymer containing 51 AAs and that is used in the treatment of diabetes [4].

AA nomenclature and the name-to-peptide method

Several years ago, the IUPAC-IUB, JCBN officially adopted a system whereby the formal and trivial chemical names of AAs are abbreviated with single letters of the English alphabet, and this system is in common use by chemists and bioscientists throughout the world (Table 1) [3]. In 2003, Wade proposed a novel method for creating biologically active peptides in which the strings of letters in personal and other names were considered as strings of IUPAC-IUB, JCBN single letter abbreviations for AAs, and used to design peptides [5-7]. The major deficiency of the name-to-peptide concept is that not all letters of the English alphabet have official IUPAC-IUB, JCBN AA assignments. Unassigned letters, or those with ambiguous assignments, are B, J, O, X, and Z. Of these five letters, O occurs most frequently (7.5%), B occurs much less frequently (1.5%), and the remaining three, J, X and Z occur with almost insignificant frequencies (0.1%) in text [8]. The name-to-peptide method was validated in 2004 by Wade, Yang, and Lea using the name of former US Secretary of State, Colin Powell, as the basis for creating the peptide, COLINPOWELL, where O represented Ornithine, a close structural analog of Lysine [9]. Upon testing, this peptide exhibited anticancer and immune stimulating properties.

The name-to-peptide method differs from standard peptide discovery methods, in that it does not use nature as a starting point, and the resulting peptide may, or may not, be found in nature. However, the

method has the ability to generate peptides of potential medical usefulness, and, due to the fact that names are important in all cultures, it also has the potential benefit of increasing interest in peptide science among the general public by facilitating a better understanding of this field of research.

Methods and Results

BLAST search for peptide, AMELIAEARHART

It would be of interest to determine if peptide, AMELIAEARHART, occurs in nature, either as an isolated peptide or as an AA segment within a larger natural protein, because such information might yield clues to potential biological functions of the peptide. These possibilities were investigated by using the peptide's AA sequence to do Basic Local Alignment Search Tool (BLAST) searches of the National Center for Biotechnology Information (NCBI) protein databases (http://blast.ncbi.nlm.nih.gov/Blast.cgi), which, at the time of the search, contained almost 11 million AA sequences [10]. BLAST search results are shown in Table 2. The AA sequence, AMELIAEARHART, was not found in its entirety, either alone or within any protein. However, numerous partial sequence matches were found for all parts of the sequence, including 164 exact matches for AMELIA, and three for EARHART. Interestingly, the sequences AMELIA and EARHART were both found in proteins of *Vitis vinifera* (grapevine), but not in the same protein. The size of the NCBI protein databases grows rapidly, as new protein AA sequences are discovered, and it is possible that in the future, an exact match may be found for the entire sequence, AMELIAEARHART. The significance of these results are that the AA sequences represented by the sequence of letters in the personal name, Amelia Earhart, are acceptable in nature, and may have useful biological properties, both in proteins and as isolated peptides.

Occurrence of portions of peptide AMELIAEARHART in the 3D structures of proteins

The BLAST program was also used to perform alignments of the peptide AA sequence with the AA sequences of proteins of known three dimensional (3D) structure in the Protein Data Bank (PDB; http://www.pdb.org/pdb/home/home.do). Two types of structures commonly found in the 3D structures of proteins are the α-helix, which has a coiled, cylindrical shape, and the β-strand, which has an extended flattened shape (Figure 3) [4]. Those portions of peptide, AMELIAEARHART, that occurred in proteins with known 3D structure were found in both α-helices and β-strands (Figure 4).

Modeling of peptide AMELIAEARHART

In addition to determining if peptide AMELIAEARHART occurs alone or in natural proteins, it also would be of interest to examine the potential structures of the name peptide in isolation as this might also provide clues to possible biological functions of the peptide. This would be of particular interest in the event that the peptide was synthesized for testing purposes.

A two dimensional (2D), wireframe model of peptide AMELIAEARHART was created with the Symyx® Draw, version 3.1.2.NET program, Symyx Technologies, Inc., and the result is shown in Figure 5. Three dimensional (3D) wireframe models were created with the Swiss PdbViewer program {v4.0.1, available from the Swiss Institute of Bioinformatics (http://spdbv.vital-it.ch/)} using omega, phi, and psi angles corresponding to a β–strand or α-helix. Each model was subjected to one round of energy minimization (5,000 steps of steepest descent), and the energy of each resulting structure was recorded. Electrostatic potential models were also made with the Swiss PdbViewer program using default parameters. The β–strand model was further modified by rotating side chains to maximize their visibility when the plane of the β–strand was displayed parallel to the plane of the computer screen. The models were then transferred to the RasWin Molecular Graphics program (Windows version 2.6-ucb) [11], and modeled as stick figures using the CPK color scheme. Molecular measurements were also made with the RasWin program. The stick figures were then transferred to the Microsoft ® Paint program (version 5.1) for the addition of letters and charges (+/-), and for overlaying electrostatic potential models upon stick figure models. The results are shown in Figures 6 and 7. Figure 8 shows that regions of opposite

(Text continued on page 10.)

Table 2. Examples of the results of a BLAST search of the NCBI protein databases (10,866,589 AA sequences) for the AA sequence, AMELIAEARHART.

Sequence Found:	Protein:	Source:[a]	Accession Number:	% Identity:
A_EL_AEARH_RT	AGAP002112-PA	*Anopheles gambiae* str. PEST	ref|XP_320926.4|	77
AMELIA_AR_AR_	mur ligase family protein	*Plesiocystis pacifica* SIR-1	ref|ZP_01912972.1|	"
AM_L_AEARH___	Nonribosomal peptide synthetase 1	*Cochliobolus heterostrophus*	gb|AAX09983.1|	61
AMELIAEA_____	Arginase	*Natranaerobius thermophilus* JW/NM-WN-LF	ref|YP_001916423.1|	"
"	Transcriptional regulator, TetR family protein	*Hoeflea phototrophica* DFL-43	ref|ZP_02168146.1|	"
AMELIAE______	Anthranilate synthase component i	*Campylobacterales bacterium* GD 1	gb|EDZ61146.1|	54
"	"	*Saccharomyces cerevisiae* YJM789	gb|EDN63066.1|	"
"	Unnamed protein product	*Vitis vinifera*	emb|CAO40359.1|	"
"	3,4-dihydroxy-2-butanone 4-phosphate synthase/GTP cyclohydrolase II	*Roseobacter* sp. SK209-2-6	ref|ZP_01754445.1|	"
___L_AEARHAR_	Adenosine deaminase	*Burkholderia dolosa* AUO158	ref|YP_002097071.1|	61
__E_I_EARHART	Unnamed protein product	*Vitis vinifera*	emb|CAO65564.1|	69
______EARHART	Purine catabolism regulator-like protein	*Streptosporangium roseum* DSM 43021	ref|ZP_04471185.1|	54
______EARHART	Hypothetical protein MICPUN_59356	*Micromonas* sp. RCC299	ref|XP_002502885.1|	"

[a]Note: *Anopheles gambiae*, mosquito that is the principal vector of malaria in Africa; *Natranaerobius thermophilus*, a halophilic, alkalithermophilic bacterium from soda lakes of the Wadi An Natrun, Egypt; *Burkholderia dolosa*, bacterium of species, proteobacteria; *Vitis vinifera*, common grape vine; *Plesiocystis pacifica*: myxobacterium isolated from Japanese coasts; *Cochliobolus heterostrophus*: fungal pathogen of corn; *Hoeflea phototrophica*: marine bacterium; *Campylobacterales bacterium*: bacterium that inhabits the gastrointestinal tract of higher animals; *Saccharomyces cerevisiae*: fungus, bakers' yeast; *Roseobacter*: marine bacterium and pathogen of oysters; *Streptosporangium roseum*: soil bacterium; *Micromonas*: photosynthetic, marine algae.

Wade Research Foundation Reports (2010) 5 (3)

Figure 3. Ribbon diagrams of two types of structure commonly found in proteins [4]: the β-strand which has a flattened, extended structure (top) and the α-helix, which has a coiled, cylindrical shape [bottom left (longitudinal view) and bottom right (cross sectional view)]. Ribbon diagrams only show the shape of the peptide backbone [i.e., successive amide bonds (-CO-NH-) within the peptide], and do not show AA side chains. Both structures shown below were modeled from a peptide containing 20 Alanine residues. In the β-strand conformation, the peptide is 66 Å long whereas in the α-helical conformation, it is only 32 Å long, a 51% reduction in length. (Note: The scale of the two objects shown is not the same, but the measured dimensions are correct)

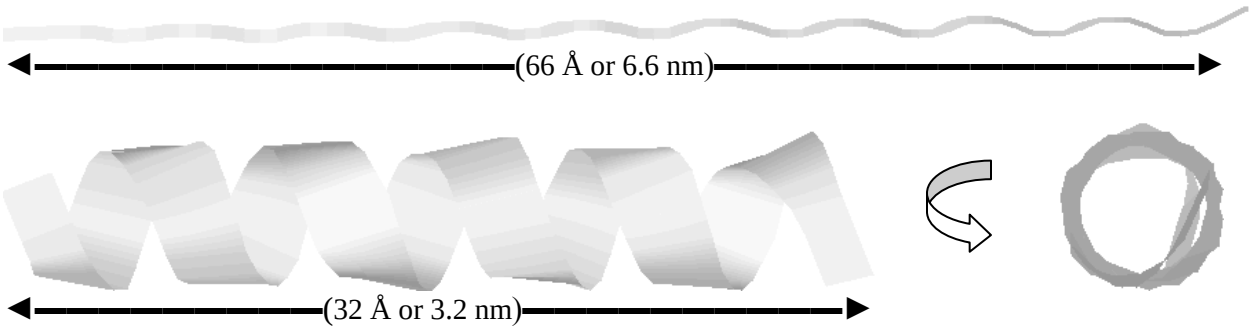
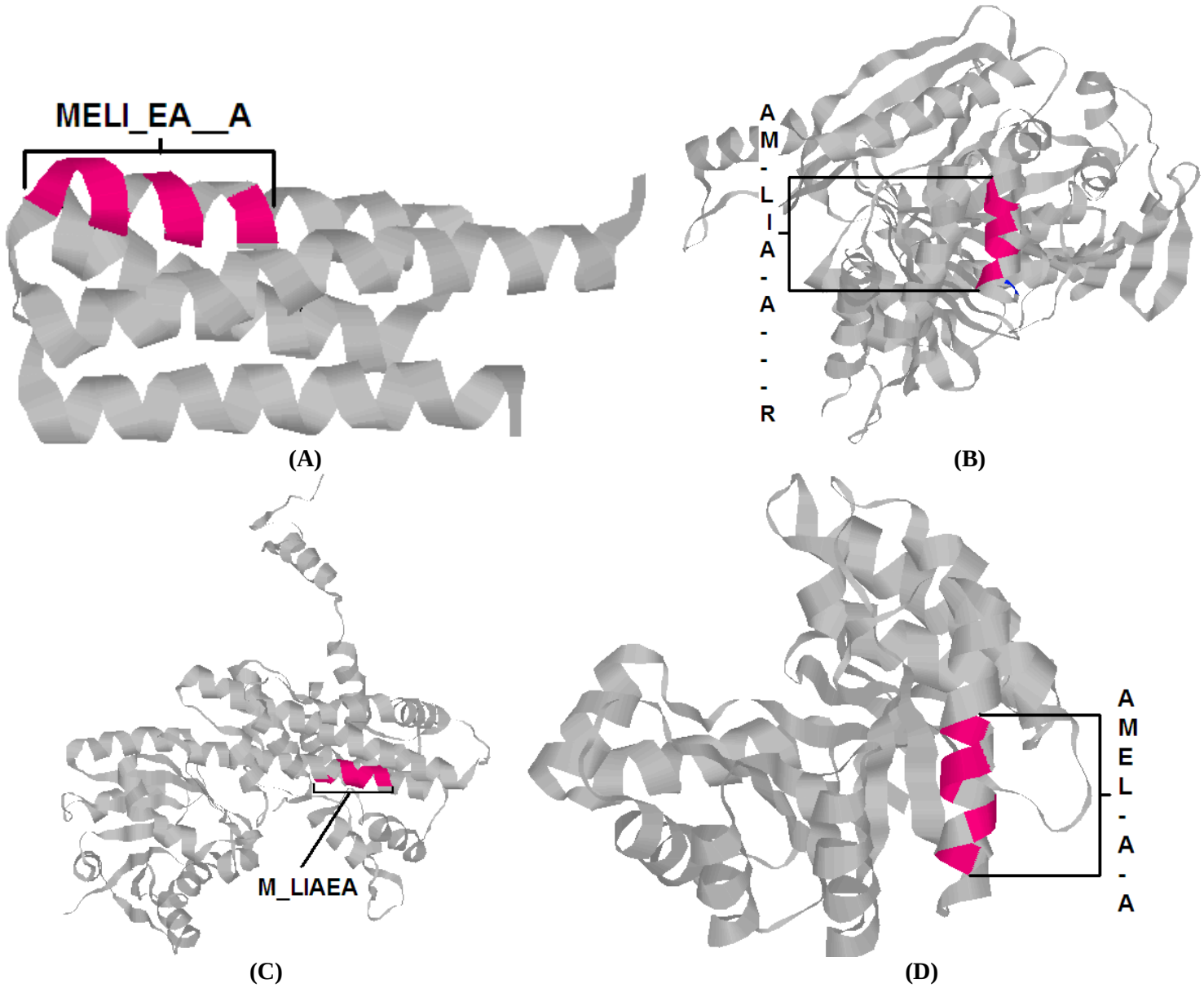
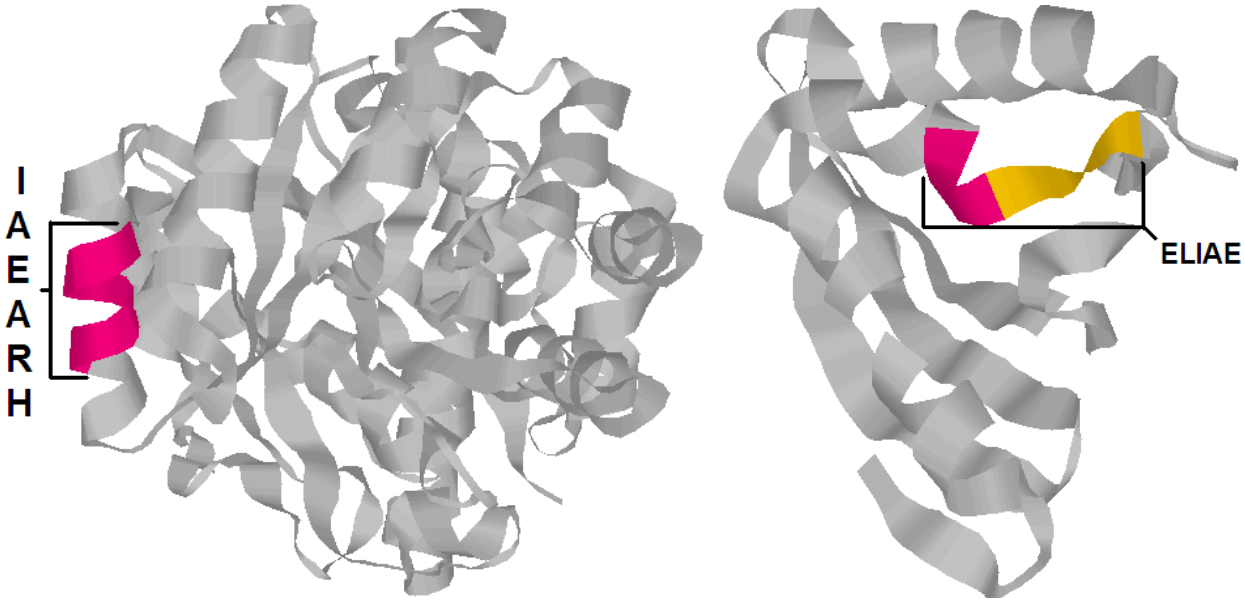


Figure 4. (See following two pages.)
(A) Ribbon diagram of the 3D structure of phosphotransferase domain of the chemotaxis protein, CHEA, from *Thermotoga maritime*, a 4 helix bundle, showing the location of sequence, MELI_EA__A, (pink color; AAs 35-43; 54% of the search sequence). The entire sequence occurs within an α-helix. Structure obtained from PDB file, 1TQG.
(B) Ribbon diagram of chain A of the protein, anthranilate synthase, from the bacterium *Salmonella typhimurium*, showing the location of the AA sequence, AM_LIA_A___R (AAs 434-445; 54% of search sequence; pink color),. All, except R445 (blue), are within an α-helix. PDB file, 1I1Q.
(C) Ribbon diagram of chain A of the protein, 6-phosphogluconate dehydrogenase, from the bacterium *Geobacillus stearothermophilus*, showing the location of the AA sequence, M_LIAEA (AAs 193-199) from AMELIAEARHART (i.e., 46% of search sequence; pink color) within an α-helix. Structure obtained from PDB file, 2W90.
(**D)** Ribbon diagram of the 3D structure of 5-methyltetrahydrofolate corrinoid/iron sulfur protein, methyltransferase, of *Moorella thermoacetica*, showing the location of the sequence, AMEL_A_A (pink color; AAs 140-147; 46% of the search sequence). The entire sequence occurs within an α-helix. PDB file, 2OGY.
(**E)** Ribbon diagram of chain A of the protein, muconate cycloisomerase I, from the bacterium, *Pseudomonas putida*, showing the location of the AA sequence, IAEARH (AAs 151-156) from AMELIAEARHART (i.e., 46% of search sequence; pink color) within an α-helix. Additional alignments that occur in the protein are the AA sequences, A--LI---A, EAR-AR. Structure from PDB file, 1F9C.
(**F)** Ribbon diagram of the 3D structure of NTF2-like protein of *Mesorhizobium loti* showing the location of sequence, ELIAE (pink and orange color, AAs 39-43; 38% of the search sequence). The pink color occurs at the end of an α-helix, and the orange color is part of a β-strand. PDB file 3FH1.
(**G)** Ribbon diagram of the 3D structure of the chromophore binding domain of the protein, bacteriophytochrome, from *Deinococcus radiodurans*, showing the location of sequence, E_IAEAR, (yellow color; AAs 185-191; 46% of the search sequence). The entire sequence occurs within a β-strand. PDB file 2O9B.

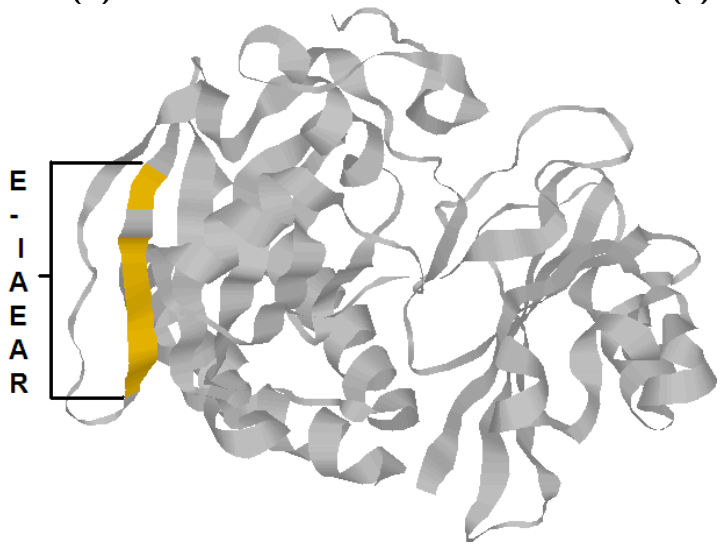


Wade Research Foundation Reports (2010) 5 (3)



(E)

(F)



(G)

Wade Research Foundation Reports (2010) 5 (3)

Figure 5. A 2D representation of the chemical structure of the peptide, AMELIAEARHART. Carbons and hydrogens are not shown. The single letter abbreviations for each AA in the peptide are shown directly under the locations of α-carbons of each AA. The average molecular mass of the peptide is 1,469, and its theoretical pI value is 6.8. Charges (+/-) on the amino (N-) and carboxyl (C-) terminal ends of the peptide, and on AA side chains, at pH 7 are shown adjacent to the N- and C-terminal ends and the side chains. The net charge on the peptide is 0.

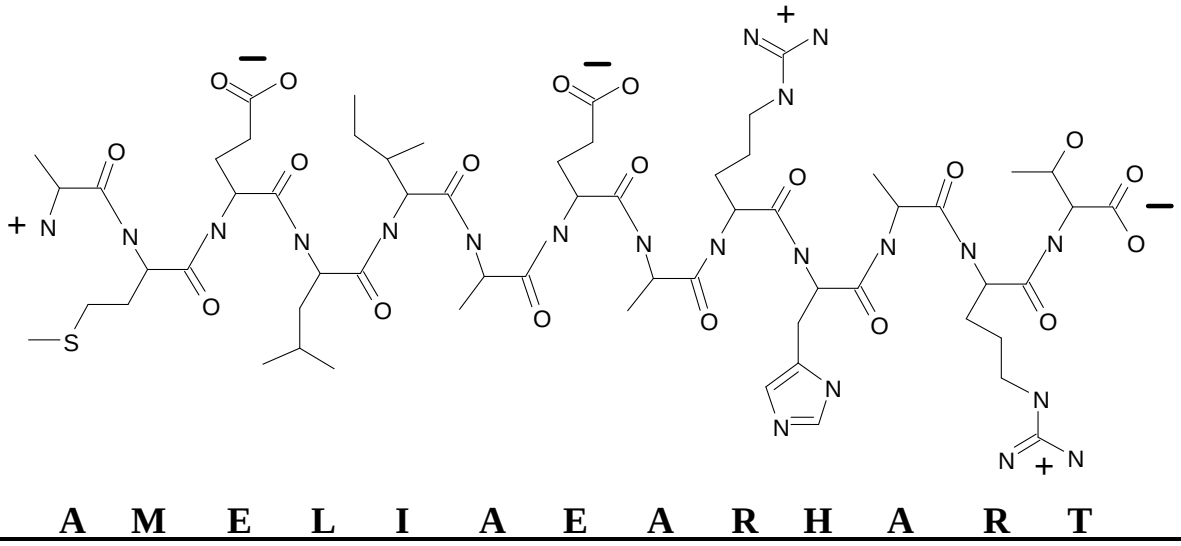
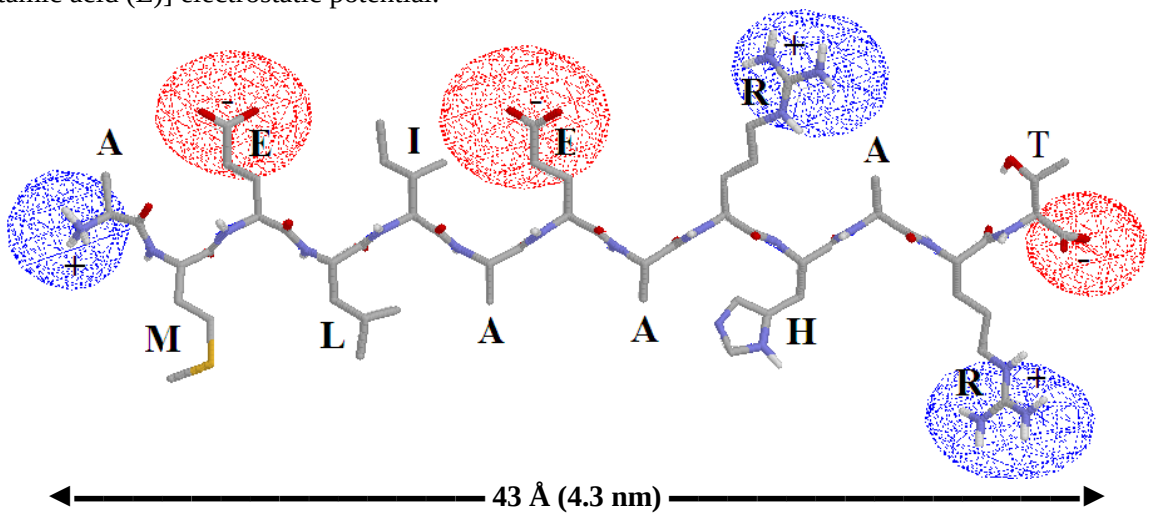
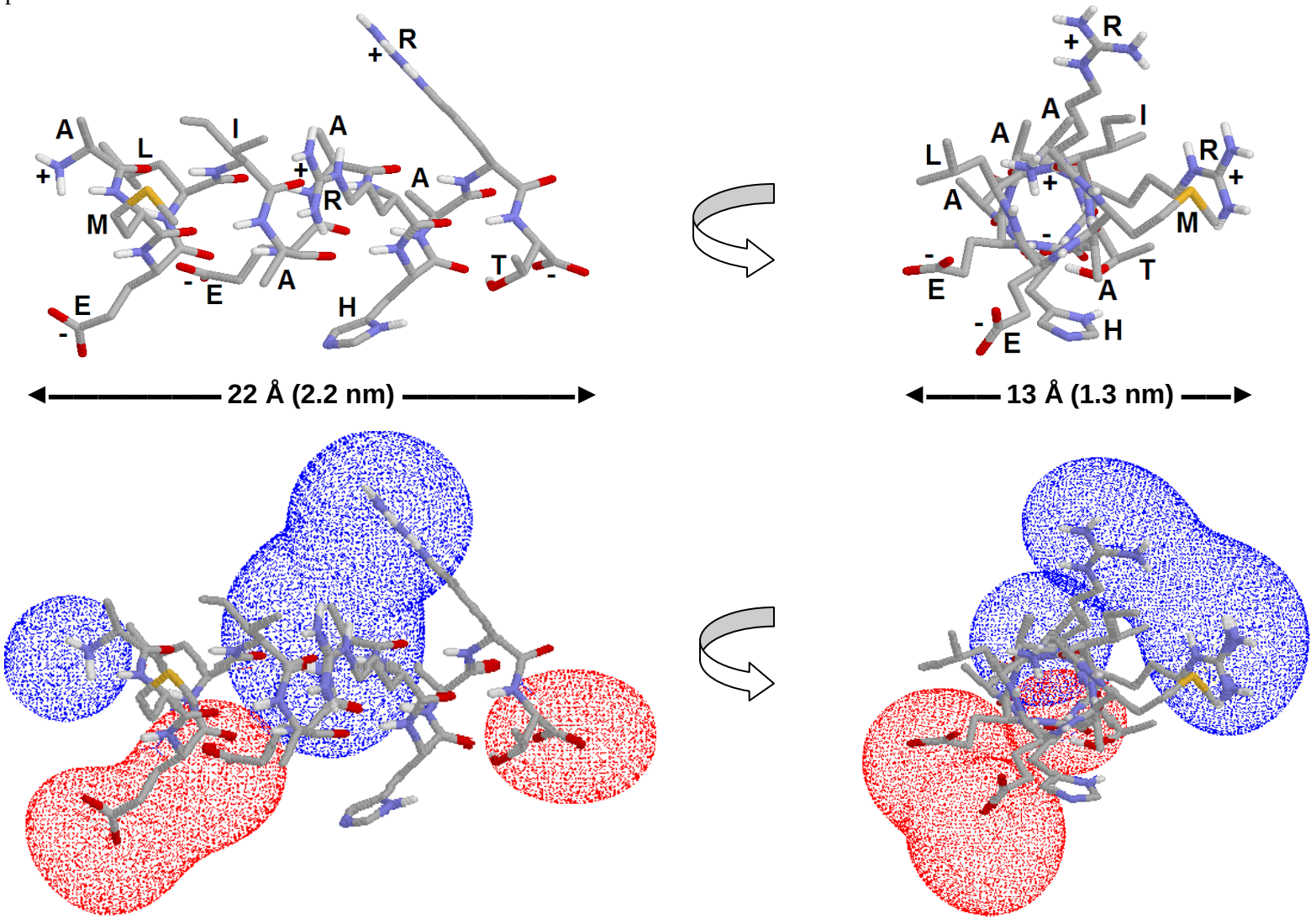


Figure 6. Combined stick figure and electrostatic potential model of peptide, **AMELIAEARHART**, as a β-strand. The amino (N-) terminal end of the peptide is on the left and the carboxyl (C-) terminal end of the peptide is on the right. The color scheme of the stick figure model is gray for carbon, blue for nitrogen, red for oxygen, yellow for sulfur, and white for hydrogen. Due to a defect in the molecular modeling program, the hydrogens of methylene (-CH_2-) and methyl (-CH_3) groups in the side chains of all AAs could not be shown. Single letter abbreviations for each AA are shown adjacent to the AAs. Charges (+/-) on the ionizable amino (N)- and carboxyl (C)-terminal ends of the peptide, on the side chain of Glutamic acid (E), and on the ionizable guanidino group in the side chain of Arginine (R) at pH 7 are shown. The net charge on the peptide would be 0. The overall dimensions of the molecule are 43 Å [4.3 nanometers (nm)] long (horizontal) and 17.5 Å wide (1.7 nm; vertical). An electrostatic potential model of the peptide was overlaid on the stick figure model, to show the regions of positive [blue colored; N-terminus and side chain of Arginine (R)] and negative [red colored, C-terminus and side chains of Glutamic acid (E)] electrostatic potential.



Wade Research Foundation Reports (2010) 5 (3)

Figure 7. Stick figure (top), and combined stick figure and electrostatic potential (bottom) models of peptide AMELIAEARHART in the α-helical configuration. The models on the left show the peptide in longitudinal view, and the models on the right show cross sectional views of the peptide. The figures on the right show that the polar and nonpolar AAs are not segregated on opposite sides of the helix, and that the peptide is not amphipathic.

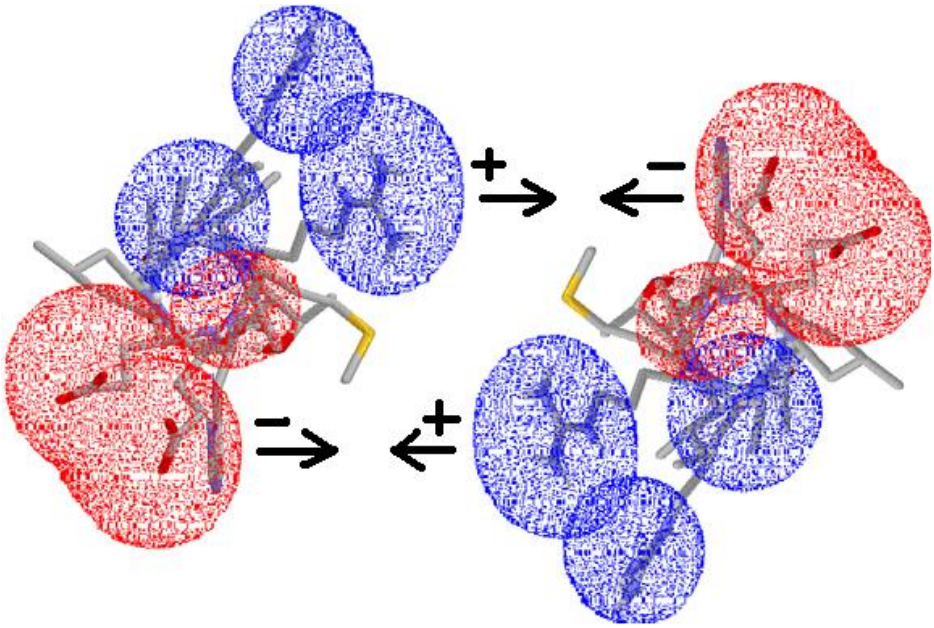


Wade Research Foundation Reports (2010) 5 (3)

Table 3. Comparison of two properties of peptide, AMELIAEARHART, in the α-helix and β-strand conformations, after energy minimization. The α-helix conformation is more compact, and more stable (lower energy) than the β-strand conformation.

Property	β-strand	α-helix	Δ
Energy (kJ/mol)	-452	-679	227
Length (Å/nm)	43 / 4.3	22 / 2.2	21 / 2.1

Figure 8. Model of the intermolecular electrostatic attractions between the charged regions of two molecules of AMELIAEARHART. Both peptides are in α-helical conformations, and are shown in end views. The model on the right was obtained by copying the model on the left, and rotating it by 180° vertically in the plane of the page. The same type of intermolecular attractions shown below would also take place between a molecule of AMELIAEARHART and any other type of charged molecule.



(Continued from page 3:)

electrostatic potential/charge in two molecules of peptide AMELIAEARHART can attract each other (+ ↔ -).

Discussion

Structures

All of the protein 3D structures shown in Figure 4, except for 2W90 [Figure 4 (C)], have their AMLEIAEARHART segments located on the surface of the protein, where they would be accessible to other molecules in the environment. Most of the protein structures of Figure 4 also have their AMELIAEARHART segments located within α-helices, but one [Figure 4 (F)] has its AMELIAEARHART segment partially in an α-helix and partially in a β-strand conformation, and another [Figure 4(G)] has its AMELIAEARHART segment entirely in a β-strand conformation. These secondary structural differences for the same AA segment illustrate the structural possibilities for peptide AMELIAEARHART, and the role of additional factors beside AA sequence (e.g., the protein environment surrounding the segment) in determining the shape of an AA sequence. In spite of these examples, it is expected that peptide AMELIAEARHART would have little, if any, structure in aqueous solutions, due to the shielding effect of water molecules on those elements of the peptide that might induce secondary structure. The peptide would only assume secondary structure (e.g., α-helix, β-strand,

or a combination of the two) when it is in environments that have lower dielectric constants than water [e.g., organic solvents, such as trifluoroethanol (TFE), lipid membranes, or solutions of high protein concentration where protein-protein binding might occur] [12], which would enable the formation of intramolecular hydrogen bonds. Table 3 indicates that the more compact α-helical conformation of the peptide would have a lower energy, and, therefore, be more stable than a more extended structure such as the β-strand. Consequently, peptide AMELIAEARHART would probably form an α-helix in environments with a low dielectric constant.

Implementing the design

Technology developed by R.B. Merrifield (1984 Nobel Prize, Chemistry), made it possible to rapidly create almost any peptide by chemical synthesis [13, 14]. This technology enables the synthesis of both naturally occurring peptides, and also peptides that do not occur in nature [12, 15]. For example, peptide AMELIAEARHART could be synthesized in substantial quantities in less than a day, and with either naturally occurring L-AAs and/or with D-AAs to make it refractory to enzymatic cleavage. Since all of the AAs in peptide AMELIAEARHART are coded for in the universal genetic code, it would also be possible to make the peptide by genetic engineering techniques [16] (Figure 9).

Properties of the hypothetical peptide

Figure 10 shows a GRand AVerage of HydropathicitY (GRAVY) plot for peptide AMELIAEARHART. It indicates that the N-terminal end of the peptide is slightly more hydrophobic than its middle or C-terminal end. This would make the N-terminal end more soluble in nonpolar environments, such as the interior of cell membranes.

Table 4 shows the sites of enzymatic cleavage of peptide AMELIAEARHART, as determined by the Expert Protein Analysis System (ExPASy) proteomics server program, Peptide Cutter (http://www.expasy.org/tools/peptidecutter/). Some of enzymes listed in the table are found in the human digestive tract {e.g., pepsins, (stomach), trypsin and chymotrypsin (pancreas), and endopeptidases (intestine)} [17]. There are nine AAs that are essential for normal growth and function in humans (i.e., they must be obtained in the diet of humans): Phenylalanine (F), Histidine (H), Isoleucine (I), Lysine (K), Leucine (L), Methionine (M), Arginine (R), Threonine (T), and Valine (V) [4, 17]. Since peptide AMELIAEARHART contains six of the nine (67%) essential AAs, the peptide could be used as a nutritional supplement.

In spite of the fact that it contains several sites for enzymatic cleavage, peptide AMELIAEARHART is predicted to be relatively stable in biological systems by the ExPASy ProtParam program (Table 5). The instability index is computed to be 24.05 indicating that the protein is stable. Replacing some or all of the L-AAs of the peptide with D-AAs [12, 15], or cyclizing the peptide [19], would reduce its susceptibility to enzymatic cleavage and prolong its biological half-life (Table 6).

Several peptides that contain multiple Arginine AAs have been found to have the ability to penetrate cell membranes, and are termed cell penetrating peptides (CPPs) [20], and at least 3 CPPs have the same or similar size as peptide AMELIAEARHART {e.g., Oligoarginine (8-9 AAs), MAP (12 AAs), Tat (13 AAs)}. There are two Arginines (R) in peptide AMELIAEARHART, located at positions 9 and 12 from the amino (N-) terminus, near the carboxyl (C-) terminal end of the peptide. Enzymatic cleavage of the C-terminal AA, Threonine (T), from the peptide would yield the peptide AMELIAEARHAR,

AMELIAEARHART ——(Enzyme)——► AMELIAEA**R**HA**R**

which contains the so-called CendR motif, R/K-X-X-R/K at the C-terminal end of the peptide, where R is Arginine, K is Lysine, and X is any AA [21-23]. The CendR motif has been found to facilitate the entry of peptides into cells via the binding of peptide to neuropilin-1 (NRP1), a transmembrane receptor on the surfaces of cells. The enzymes, tissue plasminogen activator (tPA) and urokinase-type plasminogen activator (uPA) catalyze the conversion of plasminogen to plasmin, the major enzyme responsible for

(Text continued on bottom of next page.)

Figure 9. Two major ways to synthesize peptide AMELIAEARHART. The use of solid phase peptide synthesis (SPPS) is the least complex, and may also be the most cost effective, method for producing the peptide [13, 14]. SPPS of the peptide could be done using automated procedures, and it would enable the use of non-natural D-AAs to obtain a peptide with modified properties (e.g., less susceptible to enzymatic cleavage, and, therefore, having a longer half life in biological systems) [12, 15]. Due to the small size of the peptide, it would be obtained in high yield and purity, and require a minimum of post synthetic purification. Obtaining the peptide by genetic engineering techniques would be more labor intensive, and require several complex preliminary and final purification steps than for a peptide produced by SPPS. The peptide could be produced in unlimited quantities by either method.

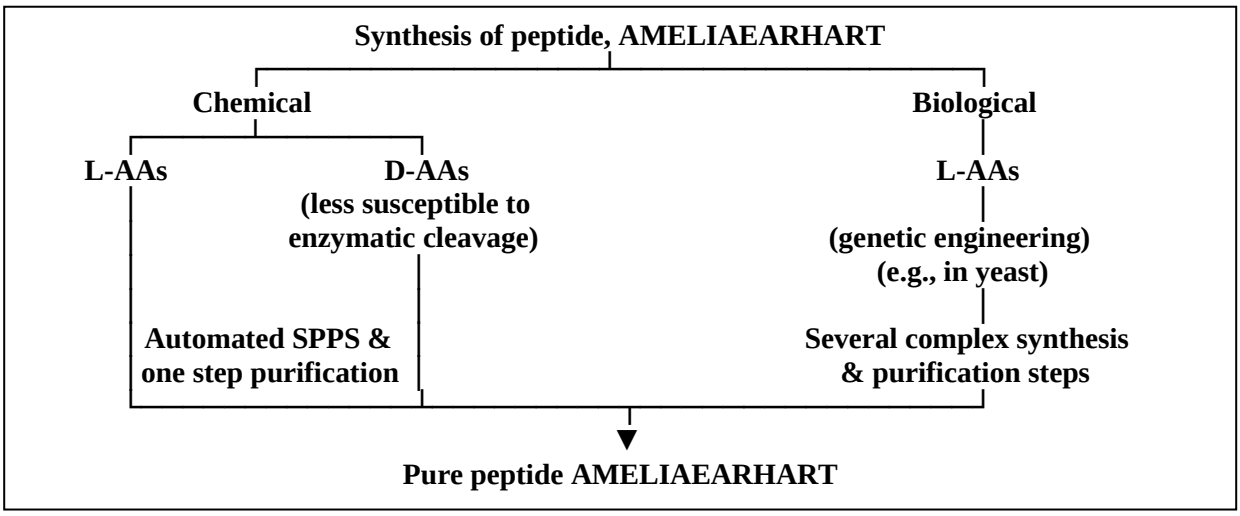
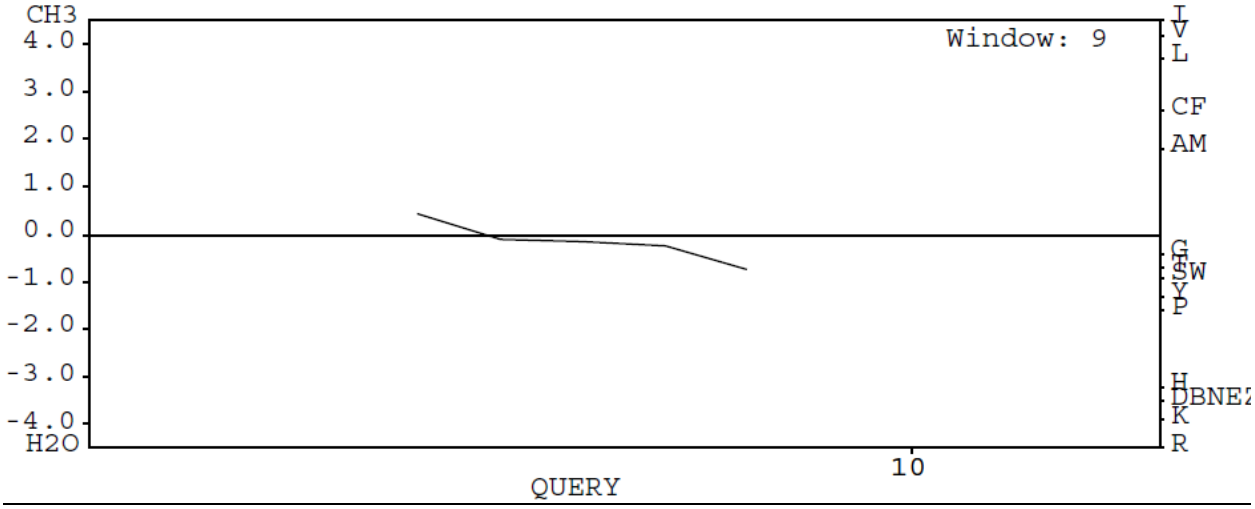


Figure 10. GRAVY plot of the hydropathicity of peptide AMELIAEARHART. The plot was prepared using an online program (http://fasta.bioch.virginia.edu/fasta_www2/fasta_www.cgi?rm=misc1).



(Text continued from previous page.)

blood clot breakdown. Both also cleave an artificial substrate, amino-4-methylcoumarin (AMC) fluorogenic peptide (i.e., R-H-A-R-AMC), that has a C-terminal AA sequence identical to the cryptic CendR motif in peptide AMELIAEARHART (Table 4) [18]. Therefore, tPA and uPA may also be capable of cleaving peptide AMELIAEARHART between Arginine 12 and Threonine 13 to expose

(Text continued on page 14.)

Table 4. Potential sites of enzymatic and chemical cleavage of peptide AMELIAEARHART as determined with the online program, Peptide Cutter from the ExPASy proteomics server (Examples 1-8), and from Reference 18 (Ex. 9). The sites of peptide cleavage are indicated by solid triangles.

Ex.:	Enzyme:	Peptide cleavage sites:
1	Arg-C proteinase or Clostripain; Trypsin	▼ ▼ A-M-E-L-I-A-E-A-R-H-A-R-T
2	Proteinase K	▼ ▼ ▼ ▼ ▼ ▼ A-M-E-L-I-A-E-A-R-H-A-R-T
3	Chymotrypsin (low specificity)	▼ ▼ ▼ A-M-E-L-I-A-E-A-R-H-A-R-T
4	Thermolysin	▼ ▼ ▼ ▼ A-M-E-L-I-A-E-A-R-H-A-R-T
5	Glutamyl endopeptidase; Staphylococcal peptidase I	▼ ▼ A-M-E-L-I-A-E-A-R-H-A-R-T
6	Asp-N endopeptidase + N-terminal Glu	▼ ▼ A-M-E-L-I-A-E-A-R-H-A-R-T
7	Pepsin (pH 1.3); Pepsin (pH >2)	▼ ▼ A-M-E-L-I-A-E-A-R-H-A-R-T
8	CNBr	▼ A-M-E-L-I-A-E-A-R-H-A-R-T
9	tPA and uPA: Artificial substrate —> (compare with —>)	▼ R-H-A-R-AMC A-M-E-L-I-A-E-A-R-H-A-R-T

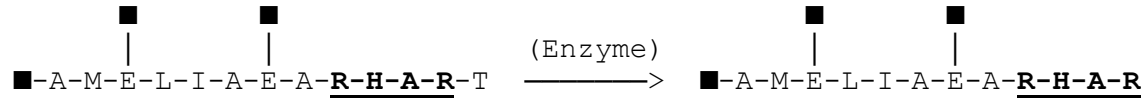
Table 5. Estimated half-life as determined with the online program, ProtParam tool, available on the ExPASy proteomics server (http://www.expasy.org/tools/protparam.html).

Organism	Estimated half-life (hours)
Mammalian reticulocytes, in vitro	4.4
Yeast, in vivo	>20
Escherichia coli, in vivo	>10

Table 6. Modifications that would increase the biological half-life of peptide AMELIAEARHART include altering the sites of enzymatic cleavage by including D-AAs (shown as lower case letters) [12, 15], or cyclizing the peptide [19].

AA Type:	Enzyme(s):	Peptide cleavage sites:
All L-AAs	Trypsin	▼ ▼ A-M-E-L-I-A-E-A-R-H-A-R-T
All L-AAs, except D-Arginine at position 9 (as lower case letter).	Trypsin	▼ A-M-E-L-I-A-E-A-**r**-H-A-R-T
All D-AAs	All	**a-m-e-l-i-a-e-a-r-h-a-r-t**
All L-AAs	All	A-M-E-L-I-A┐ │ E T-R-A-H-R-A┘

Figure 11. Use of peptide AMELIAEARHART as a carrier molecule for other molecules, such as drugs, to facilitate delivery of the second molecule into cells. AMELIAEARHART would be derivatized with the second molecule (■), for example, at the N-terminal amino (H_2N-) group, or at the side chain carboxyl (-COOH) groups of Glutamate (E) residues at positions 3 and 7. The derivatized peptide would then be introduced to the biological system where enzymatic cleavage would remove the C-terminal Threonine (T), exposing the CendR motif (in bold and underlined). The exposed CendR motif would then facilitate entry of the peptide with attached second molecule into cells via the NRP1 receptor.



(Continued from page 12.)

its cryptic CendR motif.

Peptides containing the CendR motif have also been used to facilitate delivery of other molecules (e.g., drugs) into cells, either by attachment of the other molecule to the peptide, or by coadministration of the other molecule with the peptide [21-23]. It may also be possible to use peptide AMELIAEARHART for this purpose (Figure 11). During her lifetime, Amelia Earhart was an inspiration for many, carrying their hopes and dreams as she pursued goals that led to landmark and record breaking achievements. Perhaps in the future, peptide AMELIAEARHART will carry therapeutic agents that will relieve suffering and enable others to pursue similarly ambitious goals.

References

Note: Wade Research Foundation Reports are indexed in the EBSCO Academic Search™ Complete database (http://www.ebscohost.com/titleLists/a9h-journals.htm). They are available from EBSCO, or on the website of the Wade Research Foundation (www.wade-research.com), or by sending an e-mail request to: wade-research(at)hotmail.com. Please include "Reprint Request" in the subject line.

1. Amelia Earhart: The Official Website (http://www.ameliaearhart.com/about/achievements.html).
2. Butler, S., East to the dawn: the life of Amelia Earhart, Addison-Wesley, Reading, MA, 1997, 489 p.
3. Nomenclature and symbolism for amino acids and peptides. In Biochemical Nomenclature and Related Documents, 2nd ed., C. Liébecq, Ed., Portland Press, London, UK, 1992, pp. 39-69. (http://www.chem.qmul.ac.uk/iupac/AminoAcid/)
4. Nelson, D.L., and Cox, M.M., Lehninger Principles of Biochemistry, 5th edition, W.H. Freeman and Co., New York, 2008, pp. 71-140, 682-684, 878-881.
5. Wade D. The name game: use of words composed of letters of the English alphabet as a source of novel bioactive peptides. Chemistry Preprint Archive (2003) 1: 159-170. (http://www.sciencedirect.com/preprintarchive)
6. Wade D, and Wade S. The name game: use of words composed of letters of the English alphabet as a source of novel bioactive peptides. Biopolymers Peptide Science (2003) 71: 322 (abstract P082).
7. Wade, D. The name game: use of words composed of letters of the English alphabet as a source of novel bioactive peptides, In Peptide Revolution: Genomics, Proteomics & Therapeutics, M. Chorev and T. K. Sawyer, eds., American Chemical Society, Cardiff, CA, USA, 2004, pp. 580-581.
8. Lewand, Robert (2000). Cryptological Mathematics. The Mathematical Association of America, p.36.
9. Wade, D., Yang, D., and Lea, M.A. Biological and structural properties of COLINPOWELL, a synthetic peptide amide. Wade Research Foundation Reports (2004) 1: 2-35. (http://www.wade-research.com/images/COLINPOWELL_10-25-04_.pdf)
10. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. (1997) 25: 3389-3402.

11. Sayle, R.A. and Milner, E.J., RasMol: Biomolecular graphics for all. Trends Biochem. Sci. (1995) 20: 374-376.
12. Wade, D., Boman, A., Wåhlin, B, Drain, C.M., Andreu, D., Boman, H.G., and Merrifield, R.B. All D-amino acid containing channel-forming antibiotic peptides. Proc. Natl. Acad. Sci. USA (1990) 87: 4761-65. (http://www.pnas.org/content/87/12/4761.full.pdf+html)
13. Merrifield, R.B. Solid phase peptide synthesis. I. Synthesis of a tetrapeptide. J. Am. Chem. Soc. (1963) 85 (14): 2149-2154.
14. Merrifield, R.B. Solid phase synthesis. (Nobel lecture, December 8, 1984.) In Nobel Lectures, Chemistry 1981-1990, Tore Frängsmyr, Editor-in-Charge, Bo G. Malmström, Editor, World Scientific Publishing Co., Singapore, 1992, pp. 149-175. (http://nobelprize.org/nobel_prizes/chemistry/laureates/1984/merrifield-lecture.pdf)
15. Merrifield, R.B., Wade, D., and Boman, H.G. Channel-forming antibiotic peptides containing all D-amino acids. US Patent 5585353, Issued December 17, 1996.
16. Basanta, A., Herranz, C., Gutiérrez, J., Criado, R., Hernández, P.E., and Cintas, L.M., Development of bacteriocinogenic strains of *Saccharomyces cerevisiae* heterologously expressing and secreting the leaderless enterocin L50 peptides L50A and L50B from *Enterococcus faecium* L50. Applied and Environmental Microbiology (2009) 75(8): 2382-2392.
17. Ganong, W.F. Review of Medical Physiology, 18th edition, Appleton & Lange, Stamford, Conn., 1997, pp. 275, 438.
18. Harris, J.L., Backes, B.J., Leonetti, F., Mahrus, S., Ellman, J.A., and Craik, C.S. Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries. Proc. Natl. Acad. Sci. USA (2000) 97 (14): 7754-7759.
19. Kondejewski, L.H., Farmer, S.W., Wishart, D.S., Kay, C.M., Hancock, R.E., Hodges, R.S., Modulation of structure and antibacterial and hemolytic activity by ring size in cyclic gramicidin S analogs. J. Biol. Chem. (1996) 271 (41): 25261-25268.
20. Heitz, F., Morris, M.C., and Divita, G. Twenty years of cell-penetrating peptides: from molecular mechanisms to therapeutics. Br. J. Pharmacol. (2009) 157 (2): 195-206.

21. Teesalu, T., Sugahara, K.N., Kotamraju, V.R., and Ruoslahti, E. C-end rule peptides mediate neuropilin-1-dependent cell, vascular, and tissue penetration. Proc. Natl. Acad. Sci. USA (2009) 106: 16157–16162.
22. Sugahara, K.N., Teesalu, T., Karmali, P.P., Kotamraju, V.R., Agemy, L., Girard, O.M., Hanahan, D., Mattrey, R.F., Ruoslahti, E. Tissue-penetrating delivery of compounds and nanoparticles into tumors. Cancer Cell (2009) 16: 510–520.
23. Niculescu-Duvaz, D., Springer, C., and Marais, R. C-end your drugs using peptide tags. Pigment Cell & Melanoma Res. (2010) 23(2): 157-159.

Figure 12 (left). Photograph of a plaque at Amelia Earhart Memorial Airport, Atchison, Kansas, USA.

Published July 7, 2010