

The Tetragrammaton Peptides, YHWH and YHVH

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

Abstract

This article describes a theoretical experiment that uses the International Union of Pure and Applied Chemistry's one letter system for abbreviating the chemical names of amino acids to interpret the letter sequences, YHWH and YHVH, English transliterations of the Hebrew name of God, as representing the amino acid sequences of two tetrapeptides, and proposes the chemical synthesis and biological testing of the two peptides. Both sequences of letters were found to occur commonly among the nearly 4.7 million amino acid sequences in the protein databases of the National Center for Biotechnology Information, and to be associated with proteins of known biological function. There are no barriers to the chemical synthesis of these two peptides, and they have a high probability of exhibiting biological and "other" activities.

Introduction

Among the most important biomolecules of life are proteins, polymers of amino acids (AAs) that are held together by chemical bonds, called peptide bonds [1]. They have been compared to "beads on a string", where the beads are AAs, and the beads plus string is the protein. Proteins come in a variety of sizes, ranging from polymers containing only 2 AAs to polymers containing hundreds of AAs or more. Proteins that contain less than 100 AAs are referred to as peptides (Figure 1). There are numerous proteins and peptides in the human body, where they perform functions vital for life. For example, the hormone, insulin, is a peptide containing 51 AAs that is involved in the regulation of carbohydrate and lipid metabolism, and associated with diabetes.

There are about 20 different AAs that occur naturally in proteins, and when describing the AA composition of proteins, chemists commonly use one letter abbreviations that correspond to letters of the English alphabet. These abbreviations have been officially defined by the International Union of Pure and Applied Chemistry (IUPAC)-International Union of Biochemistry and Molecular Biology, Joint Commission on Biochemical Nomenclature. They are widely used in biomedical research, and can be found in any textbook of biochemistry [2].

According to the internet encyclopedia, Wikipedia, "Yahweh is a proposed English reading of יהוה (the Tetragrammaton), the name of the God of the Jews or the people of Israel, as preserved in the original consonantal Hebrew Bible text. The four Hebrew consonants read JHWH (in German transcription) or YHVH (in English transcription). It is also common to use YHWH. Jews do not pronounce the name, but use e.g. HaShem ("The Name"), out of fear of the potential misuse of the divine name." (<http://en.wikipedia.org/wiki/Yahweh>).

Figure 1. The relationship between AAs, peptides, and proteins.

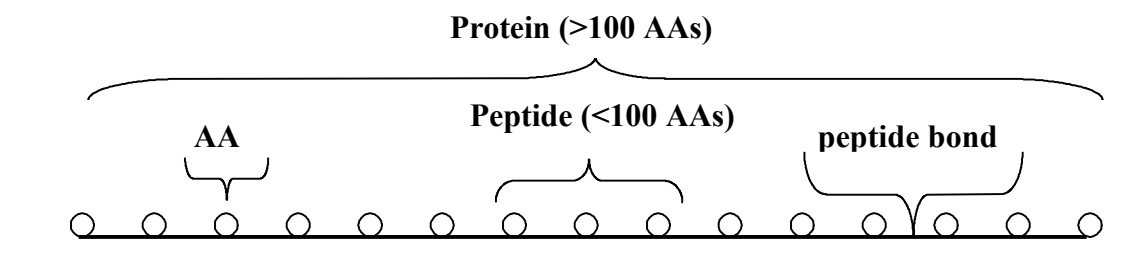
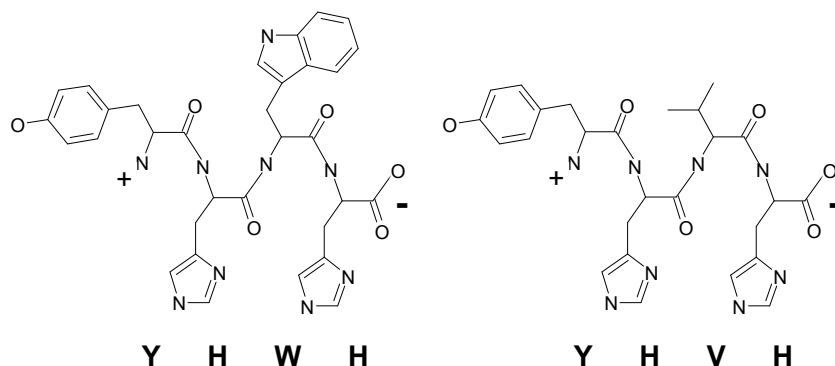


Figure 2. AA sequences of the peptides, YHWH and YHVH. Hyphens represent peptide bonds.

Tyrosine (**Y**)-Histidine (**H**)-Tryptophan (**W**)-Histidine (**H**)

Tyrosine (**Y**)-Histidine (**H**)-Valine (**V**)-Histidine (**H**)

Figure 3. Chemical structures of peptides, YHWH and YHVH. Single letter abbreviations for AAs are shown directly beneath the position of each AA in each peptide. Charges (+/-) on AAs at pH 7 are shown, and net charges on both peptides at pH 7 are 0 (neutral).



Due to technological advances developed by R.B. Merrifield (1984 Nobel Prize in Chemistry), it is possible to rapidly synthesize almost any peptide, including both naturally occurring peptides, and also peptides that do not occur in nature [3-6]. For example, it would be possible to synthesize both tetrapeptides, YHWH and YHVH, in large quantities in less than a day.

Methods

Protein database searches were done using the Basic Local Alignment Search Tool (BLAST) program for short, nearly exact matches, of the National Center for Biotechnology Information (NCBI; <http://www.ncbi.nlm.nih.gov/BLAST/>) [7]. In each case, the BLAST program was set to return the maximum number of search results (i.e., 20,000 descriptions, 20,000 alignments, and 1,000 graphic overviews).

Two dimensional models of peptides were made with the ISISTM/Draw 2.4 program (MDL Information Systems, Inc.). Three dimensional (3D) models of proteins were obtained from the RSCB Protein Data Bank (PDB; <http://www.pdb.org/pdb/home/home.do>). Figures of proteins and peptides were made using the RasWin Molecular Graphics, Windows version 2.6-ucb program (<http://mc2.cchem.berkeley.edu/Rasmol/v2.6/>) [8], and the Microsoft Paint version 5.1 program (Microsoft Corp.). Electrostatic potential diagrams were made with the Deep View/Swiss-PdbViewer v. 3.7 program (<http://www.expasy.org/spdbv/>), and the Microsoft Paint version 5.1 program.

Results and Discussion

The occurrence of YHWH and YHVH in proteins

When a new AA sequence is obtained, it is often of interest to know if the sequence occurs in natural proteins. Such information may be helpful in determining whether or not the new sequence

might have biological activities. A BLAST search of the AA sequences in the NCBI protein databases will provide such information. When BLAST searches of short, nearly exact matches, were done with YHWH and YHVH among the nearly 4.7 million AA sequences of the NCBI protein databases, 673 exact matches were returned for YHWH (Table 1), but no matches for YHVH. In fact, the BLAST program gave the error message, "No significant similarity found." When YHVH was used as the search sequence. The reason why the BLAST algorithm accepted the four letter sequence, YHWH, but not YHVH, is unknown. By simply extending the four letter sequence, YHVH, by one additional letter, G, to form the five letter sequence, YHVHG, the problem was resolved, and the search returned 95 exact matches (Table 2). This result indicates that all of the search sequences occur commonly among proteins within the NCBI databases, many of which have known biological functions. The BLAST search results could be interpreted as indicating that the YHWH and YHVH tetrapeptides might also exhibit biological activities.

Creating the Tetragrammaton peptides

There are no synthetic barriers to the creation of peptides, YHWH and YHVH, and identical AA sequences occur within many proteins. Therefore, the creation of peptides, YHWH and YHVH, is feasible, and they would have a high probability of exhibiting biological activities.

References

1. Rodwell, V.W., and Kennelly, P.J., Structures & Functions of Proteins & Enzymes: Amino Acids & Peptides, and Granner, D.K., The Diversity of the Endocrine System. in Harper's Illustrated Biochemistry, 26th Ed., R.K. Murray, D.K. Granner, P.A. Mayes, V.W. Rodwell, eds., McGraw-Hill, 2003, pp. 14-20, 449-450.
2. 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/>)
3. 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>)
4. 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 Peptide Society, San Diego, CA, USA, 2004, pp. 580-581. (Copy available upon request to author.)
5. Wade, D., Yang, D., and Lea, M.A. Biological and structural properties of COLINPOWELL, a synthetic peptide amide. Wade Res. Found. Rep. (2004) 1: 2-35. (http://www.wade-research.com/images/COLINPOWELL_10-25-04_.pdf)
6. Wade, D., Boman, A., Wählin, 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: 4761-4765. (<http://www.pnas.org/cgi/reprint/87/12/4761>)
7. Altschul, S.F., Madden, T.L., Schäffer, 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.
8. Sayle, R.A. and Milner, E.J., RasMol: Biomolecular graphics for all. Trends Biochem. Sci. (1995) 20: 374-376.

Wade Research Foundation Reports (2007) 4 (2)

Table 1. Examples of the occurrence of the four letter sequence, **YHWH**, within proteins of the NCBI protein database.

Search Sequence:	Sequence Found:	Database Accession #:	Protein and Source:	Location in protein & total AAs in protein:
YHWH	YHWH	PDB 2FJC	Antigen Tpf1 from <i>Treponema pallidum</i>	AAs 35-38; 156 AAs
“	“	PDB 1DDX	Arachidonic acid and prostaglandin bound to the cyclooxygenase active site sf Cox-2	AAs 354-357; 552 AAs
“	“	PDB 1PTH	Prostaglandin H2 synthase-1 (Cyclooxygenase I)	AAs 361-364; 576 AAs
“	“	ZP_01711907.1	N-acetyl-gamma-glutamyl-phosphate reductase [<i>Caldivirga maquilingensis</i> IC-167]	AAs 152-155; 310 AAs
“	“	XP_001387791.1	Rac GTPase-activating protein BCR/ABR [<i>Pichia stipitis</i> CBS 6054]	AAs 548-551; 710 AAs
“	“	ABN11820.1	Polyprotein [St. Louis encephalitis virus]	AAs 683-686; 3412 AAs
“	“	ABN11283.1	Prostaglandin G/H synthase 2b [<i>Oncorhynchus mykiss</i>]	AAs 375-378; 609 AAs
“	“	EAY58145.1	Galactose-1-phosphate uridylyltransferase [<i>Leptospirillum</i> sp. Group II UBA]	AAs 291-294; 334 AAs
“	“	ZP_01703888.1	DNA polymerase II, large subunit DP2 [<i>Methanococcus vannielii</i> SB]	AAs 493-496; 1131 AAs
“	“	ZP_01694000.1	Ser/Thr protein phosphatase family protein [<i>Microscilla marina</i> ATCC 23134]	AAs 84-87; 1027 AAs
“	“	<u>YP_001017672.1</u>	Glycosyl transferase, family 2 [<i>Prochlorococcus marinus</i> str. MIT 9303]	AAs 207-210; 314 AAs
“	“	XP_001301668.1	Integrase core domain containing protein [<i>Trichomonas vaginalis</i> G3]	AAs 4-7; 400 AAs
“	“	ZP_01668261.1	FAD-binding monooxygenase [<i>Anaeromyxobacter</i> sp. Fw109-5]	AAs 206-209; 407 AAs

Wade Research Foundation Reports (2007) 4 (2)

Table 2. Examples of the occurrence of the four letter sequence, **YHVH**, within proteins of the NCBI protein database.

Search Sequence:	Sequence Found:	Database Accession #:	Protein and Source:	Location in protein & total AAs in protein:
YHVH	YHVH	YP_955303.1	AMP-dependent synthetase and ligase [<i>Mycobacterium vanbaalenii</i> PYR-1]	AAs 181-185; 472 AAs
“	“	EAW90480.1	Alpha E integrin, (antigen CD103, human mucosal lymphocyte antigen 1; alpha polypeptide) [<i>Homo sapiens</i>]	AAs 556-560; 1196 AAs
“	“	ZP_01643013.1	Ppx/GppA phosphatase [<i>Stenotrophomonas maltophilia</i> R551-3]	AAs 385-389; 508 AAs
“	“	ZP_01561843.1	Glycosyl transferase, group 1 [<i>Burkholderia cenocepacia</i> MC0-3]	AAs 103-107; 414 AAs
“	“	YP_905037.1	Fatty-acid-CoA synthetase, FadD36 [<i>Mycobacterium ulcerans</i> Agy99]	AAs 183-187; 472 AAs
“	“	YP_731118.1	Dehydrogenase subunit-like protein [<i>Synechococcus</i> sp. CC9311]	AAs 145-149; 533 AAs
“	“	YP_712687.1	Esterase [<i>Frankia alni</i> ACN14a]	AAs 108-112; 343 AAs
“	“	YP_675139.1	Glycosyl transferase, group 1 [<i>Mesorhizobium</i> sp. BNC1]	AAs 98-102; 410 AAs
“	“	YP_656777.1	ORF122 [Ranid herpesvirus 1]	AAs 166-170; 438 AAs
“	“	YP_590375.1	ATP synthase F0, A subunit [<i>Acidobacteria bacterium</i> Ellin345]	AAs 140-144; 245 AAs
“	“	ZP_01306648.1	NADH dehydrogenase subunit L [<i>Oceanobacter</i> sp. RED65]	AAs 177-181; 524 AAs
“	“	ZP_01304157.1	Diaminopimelate/ornithine decarboxylase [<i>Sphingomonas</i> sp. SKA58]	AAs 106-110; 402 AAs
“	“	YP_641181.1	AMP-dependent synthetase and ligase [<i>Mycobacterium</i> sp. MCS]	AAs 180-184; 467 AAs

Figure 4. The location and 3D structure of peptide, **YHWH**, in a protein of known 3D structure, antigen Tpf1 from *Treponema pallidum* (PDB i.d. code 2FJC). The peptide backbone is shown as a ribbon and colored gray, and the **YHWH** peptide is shown in space filling format and colored blue. The two figures shown differ only by a 90° rotation about the vertical axis. The protein contains 4 cylindrical α -helices that are arranged in parallel, and the entire 4 helix bundle has a slight twist to it. The **YHWH** peptide occurs within and near the end of one α -helix.

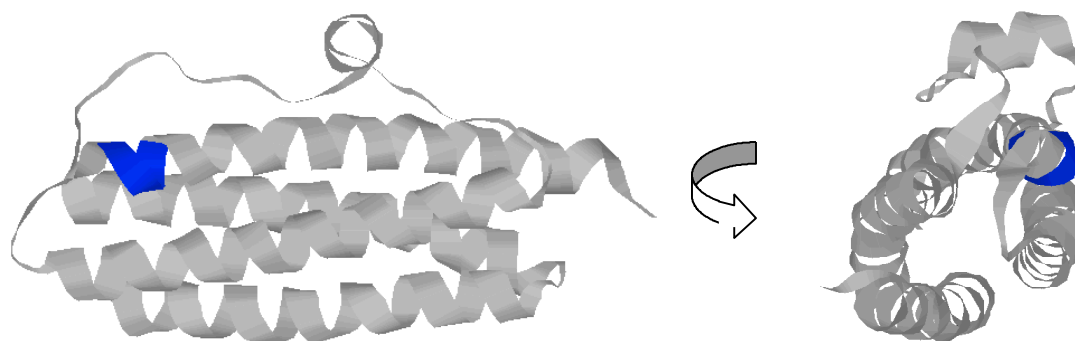


Figure 5. (Below left) A stick figure model of the isolated **YHWH** peptide as it occurs in the 3D structure of the protein of Figure 4 (above, left side, enlarged view). The position of each AA in the peptide is indicated by the adjacent single letter abbreviation for the AA. The color scheme is gray for carbon, blue for nitrogen, and red for oxygen. Hydrogens are not shown. (Below right) An electrostatic potential model of the **YHWH** peptide in the same structural orientation shown in the stick figure to the left. Blue areas are regions of positive electrostatic potential and red areas are regions of negative electrostatic potential. The electrostatic potential would have a direct effect on the ability of the peptide to interact with other molecules.

