SOFTWARE TOOL ARTICLE

PAGAL - Properties and corresponding graphics of alpha helical structures in proteins [version 3; peer review: 2 approved]

Sandeep Chakraborty¹,², Basuthkar J. Rao², Abhaya M. Dandekar¹

¹Plant Sciences Department, University of California, Davis, 95616, USA
²Department of Biological Sciences, Tata Institute of Fundamental Research, Mumbai, 400 005, India

Abstract

Alpha helices (AH) are peptide fragments characterized by regular patterns of hydrogen bonding between the carbonyl oxygen and amino nitrogen of residues regularly spaced in sequence, resulting in spiral conformations. Their preponderance in protein structures underlines their importance. Interestingly, AHs are present in most anti-microbial peptides, although they might remain in random-coil conformations depending on the solvent dielectric. For example, the cecropin component of the chimeric anti-microbial protein designed previously by our group comprises of two AHs linked by a short stretch of random coil. These anti-microbial peptides are often amphipathic (quantified by a hydrophobic moment), aligning hydrophobic residues on one surface and charged residues on the others. In the current work, we reproduce previously described computational methods to compute the hydrophobic moment of AHs - and provide open access to the source code (PAGAL). We simultaneously generated input files for TikZ (a package for creating high resolution graphics programmatically) to obtain the Edmundson wheel and showing the direction and magnitude of the hydrophobic moment, and Pymol scripts to generate color coded protein surfaces. Additionally, we have observed an empirical structural property of AHs: the distance between the Cα atoms of the ith and (i+4)th residue is equal to the distance between the carbonyl oxygens of the ith and (i+4)th residue. We validated this using 100 non-homologous high resolution structures from the PISCES database. The source code and manual is available at http://github.com/sanchak/pagal and on http://dx.doi.org/10.5281/zenodo.11136.

Keywords

Anti-microbial peptides, protein, computational methods

Open Peer Review

Invited Reviewers

1

2

Reviewer Status

Invited Reviewers

1

2

Second version: 19 Sep 2014, 3:206 (https://doi.org/10.12688/f1000research.4952.2)

1 Ramakrishnan Nagaraj, Centre for Cellular and Molecular Biology, Hyderabad, India
2 Guangshun Wang, University of Nebraska Medical Center, Omaha, USA

Any reports and responses or comments on the article can be found at the end of the article.
Corresponding author: Sandeep Chakraborty (sanchak@gmail.com)

Competing interests: No competing interests were disclosed.

Grant information: AMD wishes to acknowledge grant support from the California Department of Food and Agriculture PD/GWSS Board. BJ acknowledges financial support from Tata Institute of Fundamental Research (Department of Atomic Energy). Additionally, BJR is thankful to the Department of Science and Technology for the JC Bose Award Grant.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2015 Chakraborty S et al. This is an open access article distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the Creative Commons Zero "No rights reserved" data waiver (CC0 1.0 Public domain dedication).

How to cite this article: Chakraborty S, Rao BJ and Dandekar AM. PAGAL - Properties and corresponding graphics of alpha helical structures in proteins [version 3; peer review: 2 approved] F1000Research 2015, 3:206 (https://doi.org/10.12688/f1000research.4952.3)

Introduction
A protein structure is formed by well ordered local segments defined by the hydrogen-bonding pattern of the peptide backbone (secondary structures), and conformations that lack any regular arrangement (random coils). The most prevalent secondary structures are alpha helices (AH) and β sheets, while other conformations like π-helix occur rarely in natural proteins. AHs are right-handed spiral conformations which have a hydrogen bond between the carbonyl oxygen (C=O) of every residue and the alpha-amino nitrogen (N-H) of the fourth residue away from the N-terminal.

DSSP is the official program used to assign secondary structure to a protein when the atomic coordinates are known.s3 Several methods can also predict an AH from the sequence.s4,5 Essentially, any structure prediction tool can be used to predict an AH from the sequence by first predicting the structure and then applying DSSP to the predicted structure.s6-8

The niche of AHs in protein structures is widespread. AHs are the functionally significant element in several motifs (DNA binding motifs), and the key components of any protein that permeates biological membranes.s9 AHs are also almost always present in anti-microbial peptides (AMP)s1, although they may remain in random-coil conformations depending on the solvent dielectric.s12,13 For example, it has been recently shown that certain peptides are in random coil conformations, and achieve helical structures only by interacting with the anionic membrane model that has the same head group as the major anionic phosphatidylglycerols in bacterial membranes.s14 For example, cecropin B, a component of a chimeric protein with anti-microbial properties that provides grapevines with enhanced resistance against the Gram-negative pathogen Xylella fastidiosa.s15, is composed of two AHs connected by a small random coil.s16 Other AMPs comprise only a single AH.s17,18 These peptides are characterized by a strong hydrophobic surface (defined by a hydrophobic moment)s19, and often have charged residues, either anionic or cationic, aligned on the opposite surface.s20 Previously, Jones et al. have implemented computational methods to extract the characteristics of AHs.s21

In the current work, we first observe and propose an empirical structural property of AHs: that the distance between the Cα atoms of the i-th and (i+4)th residue (denoted by D(Cαi/Cαi+4)) is (almost) equal to the distance between the carbonyl oxygens of the i-th and (i+4)th residue (D(Oi/Oi+4)). We validate our hypothesis on a set of 100 high resolution, non-homologous proteins (which have 775 AHs) taken from the PISCES database (http://dunbrack.fccc.edu/PISCES.php). Figure 1 shows the plot of the difference between D(Cαi/Cαi+4) and D(Oi/Oi+4) for AHs specified in the PDB files (in red, mean=0.16 Å, standard deviation (sd)= 0.34 Å), and for all residues separated by four residues but not part of a helix (in blue, mean=0.71 Å, sd=0.75 Å).

Results and discussion
Validation of empirical property
We have observed an empirical structural property that applies to the residues of any AH: the distance between the Cα atoms of the i-th and (i+4)th residue is equal to the distance between the carbonyl oxygens of the i-th and (i+4)th residue. This hypothesis is validated on a set of high resolution non-homologous 100 proteins (775 AHs) taken from the PISCES database. Next, we implement the methodologies described previously to compute the hydrophobic moments for AHs using the hydrophobicity scale used in PAGAL - Properties and corresponding graphics of alpha helical structures in proteins. The current work is based on peptides that have solved structures which satisfy the AH property. In reality, due to conformational changes depending on solvent properties, the hydrophobic moment is not unique. There are other programs available online to do similar processing (http://rrzlab.ucr.edu/scripts/wheel/ for example). We also specify a metric associated with each helix - the ratio of the positive to the negative residues (RPNR) in the AH - which helps identify AHs with a particular kind of charge distribution on their surface. The results are outputted as the input to a graphical program TikZ (for the Edmundson wheel and hydrophobic moment), and Pymol scripts (for showing the peptide surface). The source code and manual available at http://github.com/sanchak/pagal and on http://dx.doi.org/10.5281/zenodo.11136.
Algorithm 1. Calculate hydrophobic moment

Input: αH: α helix - either PDB or fasta sequence
Input: TableHS: Hydrophobic scale
Output: TikZIN: TikZ input file
Output: PymolIN: Pymol input file

begin
Radius = 5 ; // Radius of Edmundson wheel
initangle = 90 ; // first residue is at 12 o’clock..
loopcnt = 0 ;
finalvechydro = undefined ;
centre = (0,0);
foreach Residue i in αH do
    /* Find X,Y coordinate on the Edmundson wheel */
    angle = initangle - loopcnt * 100 ;
    x = Radius * cos(val) ;
    y = Radius * sin(val) ;
    thispoint = (x,y);
    /* Get Hydrophobic moment */
    vector = MakeVectorFrom2Points(centre,thispoint) ;
    hydrophobicvalue = GetHydrophobicScaleForResidue(TableHS, Residue i) ;
    tmpvec = normal(vector) * hydrophobicvalue ;
    finalvechydro = finalvechydro is not defined? tmpvec : finalvechydro + tmpvec;
    loopcnt++ ;
end
WriteTikzScript();
WritePymolScript();
end

Figure 1. Plot of the difference between D(Cαi/Cαi+4) and D(Oi/Oi+4). All 775 AHs specified in the PDB files from the 100 non-homologous high resolution structures taken from the PISCES database are in red (mean=0.16 Å, standard deviation (sd)=0.34 Å). All residues separated by four residues but not part of a helix are in blue (mean=0.71 Å, sd=0.75 Å). All AHs specified in the PDB files after correction are in green (mean=0.095 Å and sd=0.14 Å).

The Pymol helix identification program shows part of this stretch as a random coil (Lys178 in Figure 2-a). Moreover, the distance between the carbonyl oxygen (C=O) and the alpha-amino nitrogen (N-H) of the fourth residue away from the N-terminal is 7.6 Å, which makes it improbable for them to have a hydrogen bond, the primary requisite to be part of an AH. The D(Cαi/Cαi+4) and D(Oi/Oi+4) for this pair is 9 Å and 8 Å, respectively: a difference of 1 Å. Even in cases where the distance between C=O and N-H is within the 3.6 Å typically required for a hydrogen bond (PDBid: 1ELU, 12th helix), the distances D(Cαi/Cαi+4) and D(Oi/Oi+4) for the residue pair His292-Gly296 is 6.9 Å and 3.4 Å, respectively: a difference of 3.4 Å (Figure 2b). In short, the helix annotation in the PDB database is often incorrect. Removing these problematic residues reduces the mean distance to 0.095 Å and the sd to 0.14 Å (Figure 1).

There is variation in the D(Cαi/Cαi+4) even when considering the same pair of residues. For example, taking all pairs of Arg and Lys in the 775 AHs analyzed (Table 1), we see that the values can vary from 6.5 Å in PDBid:1H16 (helix26, pair Arg583-Lys587) to 5.8 Å in PDBid:1EYH (helix5, pair Arg72-Lys76). However, as hypothesized, D(Oi/Oi+4) is the same as D(Cαi/Cαi+4).
Edmundson wheel and the hydrophobic moment
The Edmundson wheel\(^1\) has been the standard way of visualizing AHs for a long time now, although there are other methods (Wenxiang diagram\(^2\)) to represent AHs. The Edmundson wheel shows the alignment of residues as one looks through the helix, and gives an approximate idea of the various properties of the AH. For example, a color coding differentiation of the polar and non-polar residues gives an approximation of the hydrophobic propensity of the AH. A more mathematical representation of the hydrophobic propensity is to represent each residue with a value and a sign (direction). This results in a vector representation, called the hydrophobic moment\(^3\). We have chosen the hydrophobic scale from\(^4\) (Table 2), although any other hydrophobic scale could be also used. The color coding is as follows: all hydrophobic residues (positive values in Table 2) are colored red, while hydrophilic residues (negative values in Table 2) are colored in blue: dark blue for positively charged residues, medium blue for negatively charged residues and light blue for amides. We now show the PAGAL representation of a few AH peptides.
Cecropin. A synergistic combination of two critical immune functions, pathogen surface recognition and lysis, resulted in a chimeric protein with anti-microbial properties against the Gram-negative Xylella fastidiosa. The lytic domain is cecropin B, which attacks conserved lipid moieties and creates pores in the X. fastidiosa outer membrane. Cecropin B consists of two AHs, joined by a short stretch of random coil. Figure 3a and b shows the Edmundson wheel and hydrophobic moment of the two AHs. It can be seen that the N-Terminal AH has a large hydrophobic moment, as well as a specific positive charge distribution. The hydrophobicity of this amphipathic AH has significant bearing on the anti-microbial properties of the peptide. This can also be seen in a Pymol rendering of the peptide surface (Figure 4). The Pymol script for this rendering is automatically generated by PAGAL. On the other hand, the C-Terminal AH comprises mostly of hydrophobic residues. Cecropin-like peptides use the synergy of these two helices - the N-terminal attaches to charged ion on the membrane, and the hydrophobic C-terminal permeates the hydrophobic inter-membrane region (known as the ‘carpet’ model).

Cathelicidin LL-37. Cathelicidin LL-37 is a critical component of the innate human immune system that protects humans against infectious diseases by targeting anionic phosphatidylglycerols in the pathogenic bacterial membranes.

**Figure 3.** Visualizing the Edmundson wheel and hydrophobic moment of some alpha helices. All hydrophobic residues are colored in red, while hydrophilic residues are colored in blue: dark blue for positively charged residues, medium blue for negatively charged residues, and light blue for amides. (a) N-Terminal helix of cecropin B. (b) C-Terminal helix of cecropin B. (c) KR-12 peptide fragment from cathelicidin LL-37. (d) De novo designed peptide (SP1-1) with anti-microbial activity.
Recent work has demonstrated a 12-residue peptide (KR-12) corresponding to residues 18 to 29 of LL-37 is toxic to bacterial, but not human cells. Figure 3c shows the Edmundson wheel and hydrophobic moment of KR-12. The demarcation of the polar and non-polar residues is quite evident. The predominance of positively charged residues in the polar side of the peptide is also clearly visible.

**De novo designed AMPs for plant protection.** The de novo design of small AMPs that inhibit plant pathogens was the focus of a recent work. One of the most promising candidates was a small peptide (SP1-1 - RKKRLKLLKRL, Figure 3d), which was “highly active against a broad spectrum of bacteria, but showed low hemolytic activity”. Although the hydrophobic moment of this peptide is much smaller than that of KR-12 (Figure 3c), possibly due to the presence of Arg4 on the hydrophobic surface, the distribution of positively charged residues in this peptide is greater than for KR-12.

**Ratio of the positive to the negative residues (RPNR).** Often, it is desirable to choose a large distribution of charged residues of a certain kind (anionic or cationic) on the hydrophilic surface. One possible method for quantifying this would be to compute a ‘charge moment’, similar to the computation of hydrophobic moments. However, such an evaluation would determine certain clearly distributions to be the same. For example, assume one semicircle of the wheel comprised only positive residues, and the other hydrophobic residues (Figure 5a). This is a slightly modified version of KR-12 from cathelicidin LL-37. If one positive residue (R5) were moved from the hydrophilic side to the hydrophobic side (I7) and replaced with a negative residue (D7) (Figure 5b), the ‘charge moment’ would remain the same, although the two conformations are clearly not the same. Note that the hydrophobic moment is also different, as expected. Therefore, the ‘charge moment’ is not an accurate metric. This is underlined by the fact that replacing a hydrophilic...
serine on the hydrophobic face with a hydrophobic residue (Ala or Val) enhanced the antimicrobial peptide activity in LL-23, a natural peptide derived from the N-terminal of LL-37\(^8\). Thus, we resort to a simple metric to allow one to choose peptides with a large proportion of charged residue of a single kind: the ratio of the positive to the negative residues (RPNR). The two peptides mentioned above will have different RPNRs: 1 (Figure 5a) and 0.85 (Figure 5b). Also, the current method is unable to discriminate the possible effects of substituting similar amino acids (for example replacing an arginine by a lysine). These effects are complex and difficult to computationally model, for the consequences of the substitution of arginines for lysines is also modulated by the nature of the peptide into which the substitution is made\(^14\). Such substitutions (applied to β-defensins also, and not AH peptides) also hold promise as future therapeutic drugs\(^29\).

Output formats
PAGAL generates a TiKZ input file for drawing the Edmundson wheel and showing the hydrophobic moment (Supplementary File TikZInput.doc). TiKZ is a package “for creating graphics programmatically” - http://www.texample.net/tikz/. PAGAL also generates a Pymol script to the peptide structure using the same color coding used in for the Edmundson wheel (Supplementary File PymolInput.doc).

Software availability
Latest source code
http://github.com/sanchak/pagal

References


Open Peer Review

Current Peer Review Status: ✓ ✓

Version 2

Reviewer Report 13 October 2014

https://doi.org/10.5256/f1000research.5695.r6394

© 2014 Nagaraj R. This is an open access peer review report distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

 ✓ Ramakrishnan Nagaraj
  Centre for Cellular and Molecular Biology, Hyderabad, India

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 27 Oct 2014

Sandeep Chakraborty, University of California, Davis, USA

We would like to thank you once again for reviewing our paper. We would also like bring to your notice an application of the PAGAL program in characterizing the proteome of the Ebola virus - http://f1000research.com/articles/3-251/v1

Competing Interests: No competing interests were disclosed.

Reviewer Report 29 September 2014

https://doi.org/10.5256/f1000research.5695.r6250

© 2014 Wang G. This is an open access peer review report distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

 ✓ Guangshun Wang
  Department of Pathology and Microbiology, University of Nebraska, Omaha, NE, USA
**Competing Interests:** No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

---

Author Response 27 Oct 2014

**Sandeep Chakraborty**, University of California, Davis, USA

We would like to thank you once again for reviewing our paper. We would also like bring to your notice an application of the PAGAL program in characterizing the proteome of the Ebola virus - [http://f1000research.com/articles/3-251/v1](http://f1000research.com/articles/3-251/v1).

**Competing Interests:** No competing interests were disclosed.

---

**Version 1**

Reviewer Report 12 September 2014

https://doi.org/10.5256/f1000research.5288.r6102

© 2014 **Wang G**. This is an open access peer review report distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Guangshun Wang**
Department of Pathology and Microbiology, University of Nebraska, Omaha, NE, USA

The authors aimed at further define features for helices, which are key structural elements in polypeptides. Based on the traditional hydrophobic moment, they proposed the concept of charge moment. It is interesting but I am a little bit of doubtful how useful the charge moment will be. In particular, the authors found the “Swapping one positive residue (R5) from the hydrophilic side with I7 and replacing it with a negative residue (D7), results in the same ‘charge moment’” (Figure 5). In the case of helical antimicrobial peptides, such a swap may have a detrimental effect on peptide antimicrobial activity. For example, even a change of a hydrophilic residue serine on the hydrophobic face with a hydrophobic residue influenced peptide activity ([Wang G et al., 2012](http://f1000research.com/articles/3-251/v1)). The authors may refine this idea and conceive the possible use of charge moment. Based on charge moment, is there any clue that interfacial charged residues of antimicrobial peptides play a larger role than non-interfacial ones in determining antimicrobial activity? Will it be possible to incorporate the observation that arginines are usually more important than lysines in determining peptide activity? (see **Mishra B et al., 2013**)?

**Competing Interests:** No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.
Author Response 17 Sep 2014

Sandeep Chakraborty, University of California, Davis, USA

We would like to thank you for taking the time and reviewing our paper. Please find our response to your suggestions below. We have incorporated the changes in the new version.

The authors aimed at further define features for helices, which are key structural elements in polypeptides. Based on the traditional hydrophobic moment, they proposed the concept of charge moment. It is interesting but I am a little bit of doubtful how useful the charge moment will be. In particular, the authors found the “Swapping one positive residue (R5) from the hydrophilic side with I7 and replacing it with a negative residue (D7), results in the same ‘charge moment’” (Figure 5). In the case of helical antimicrobial peptides, such a swap may have a detrimental effect on peptide antimicrobial activity. For example, even a change of a hydrophilic residue serine on the hydrophobic face with a hydrophobic residue influenced peptide activity (Wang G et al., 2012). The authors may refine this idea and conceive the possible use of charge moment. Based on charge moment, is there any clue that interfacial charged residues of antimicrobial peptides play a larger role than non-interfacial ones in determining antimicrobial activity? Will it be possible to incorporate the observation that arginines are usually more important than lysines in determining peptide activity (see Mishra B et al., 2013)?

The relevant reference you have pointed out (Wang G et al., 2012) underlines the point are making - that the charge moment (analogous to the hydrophobic moment) is not a good metric. Thus, `we resort to a simple metric to allow one to choose peptides with a large proportion of charged residue of a single kind: the ratio of the positive to the negative residues (RPNR)`. I apologize that this point has been confusing, so I have explicitly clarified this aspect. Also, the reference helps to establish our point.

Unfortunately, we would be unable to differentiate between lysines and arginines using the current methodology. However, this is a salient point that we need to discuss so that future work may address this.

We hope to have addressed your concerns with these modifications.

Competing Interests: No competing interests were disclosed.

Reviewer Report 12 September 2014

https://doi.org/10.5256/f1000research.5288.r5988

© 2014 Nagaraj R. This is an open access peer review report distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ramakrishnan Nagaraj
Centre for Cellular and Molecular Biology, Hyderabad, India
1. AHs are not invariably present in ALL antimicrobial peptides. The formation of AHs in antimicrobial peptides is dependent on solvent conditions. The statement to this effect should be edited appropriately.

2. The proposed empirical structural property of AHs is OK as it is validated on a set of high resolution non-homologous proteins from a database.

3. Extension of the analysis to antibacterial peptides such as cecropin and LL37 may not be valid. These peptides are unstructured in water and fold into amphipathic helical structure only in media of low dielectric constant. Also, the helical structures are not rigid as observed in protein crystal structures and show considerable conformational flexibility in solution. Hence, a unique value of hydrophobic moment will not be very meaningful. Solvent effects need to be taken into account. This point (3) needs to be addressed and clarified.

**Competing Interests:** No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 17 Sep 2014

Sandeep Chakraborty, University of California, Davis, USA

We would like to thank you for taking the time and reviewing our paper, and for your helpful comments. We have modified our manuscript to reflect this changes, and hope that these are satisfactory.

**Competing Interests:** No competing interests were disclosed.

---

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com