

RESEARCH ARTICLE

# High-Resolution Crystal Structures of Protein Helices Reconciled with Three-Centered Hydrogen Bonds and Multipole Electrostatics

Daniel J. Kuster<sup>1</sup>, Chengyu Liu<sup>2</sup>, Zheng Fang<sup>1</sup>, Jay W. Ponder<sup>2</sup>, Garland R. Marshall<sup>3\*</sup>

**1** Department of Biomedical Engineering, Washington University, St. Louis, MO, United States of America, **2** Department of Chemistry, Washington University, St. Louis, MO, United States of America, **3** Department of Biochemistry and Molecular Biophysics, Washington University, St. Louis, MO, United States of America

\* [garlandm@gmail.com](mailto:garlandm@gmail.com)



**OPEN ACCESS**

**Citation:** Kuster DJ, Liu C, Fang Z, Ponder JW, Marshall GR (2015) High-Resolution Crystal Structures of Protein Helices Reconciled with Three-Centered Hydrogen Bonds and Multipole Electrostatics. PLoS ONE 10(4): e0123146. doi:10.1371/journal.pone.0123146

**Academic Editor:** Freddie Salsbury, Jr, Wake Forest University, UNITED STATES

**Received:** October 24, 2014

**Accepted:** February 24, 2015

**Published:** April 20, 2015

**Copyright:** © 2015 Kuster et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper.

**Funding:** GRM acknowledges support from National Institutes of Health (NIH) (R01-GM068460) and the Department of Defense (HDTRA1-09-CHEM-BIO-BAA). DJK acknowledges fellowship support from the Ewing Marion Kauffman Foundation and the Department of Biomedical Engineering at Washington University. JWP acknowledges support from the NSF (CHE-0535675) and NIH (R01-GM6955302). The funders had no role in study design, data collection

## Abstract

Theoretical and experimental evidence for non-linear hydrogen bonds in protein helices is ubiquitous. In particular, amide three-centered hydrogen bonds are common features of helices in high-resolution crystal structures of proteins. These high-resolution structures (1.0 to 1.5 Å nominal crystallographic resolution) position backbone atoms without significant bias from modeling constraints and identify  $\Phi = -62^\circ$ ,  $\psi = -43^\circ$  as the consensus backbone torsional angles of protein helices. These torsional angles preserve the atomic positions of  $\alpha$ - $\beta$  carbons of the classic Pauling  $\alpha$ -helix while allowing the amide carbonyls to form bifurcated hydrogen bonds as first suggested by Némethy et al. in 1967. Molecular dynamics simulations of a capped 12-residue oligoalanine in water with AMOEBA (Atomic Multipole Optimized Energetics for Biomolecular Applications), a second-generation force field that includes multipole electrostatics and polarizability, reproduces the experimentally observed high-resolution helical conformation and correctly reorients the amide-bond carbonyls into bifurcated hydrogen bonds. This simple modification of backbone torsional angles reconciles experimental and theoretical views to provide a unified view of amide three-centered hydrogen bonds as crucial components of protein helices. The reason why they have been overlooked by structural biologists depends on the small crankshaft-like changes in orientation of the amide bond that allows maintenance of the overall helical parameters (helix pitch ( $p$ ) and residues per turn ( $n$ )). The Pauling  $3.6_{13}$   $\alpha$ -helix fits the high-resolution experimental data with the minor exception of the amide-carbonyl electron density, but the previously associated backbone torsional angles ( $\Phi$ ,  $\Psi$ ) needed slight modification to be reconciled with three-atom centered H-bonds and multipole electrostatics. Thus, a new standard helix, the  $3.6_{13/10^-}$ , Némethy- or N-helix, is proposed. Due to the use of constraints from monopole force fields and assumed secondary structures used in low-resolution refinement of electron density of proteins, such structures in the PDB often show linear hydrogen bonding.

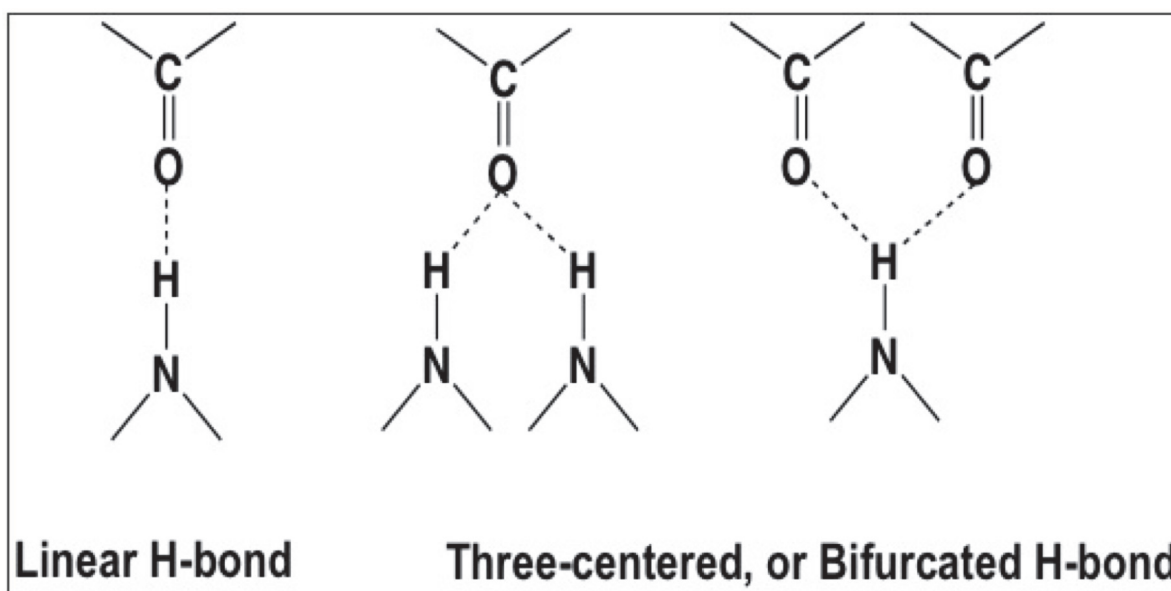
and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Introduction—Fundamental Assumptions of the Classic $\alpha$ - and $3_{10}$ -Helical Models

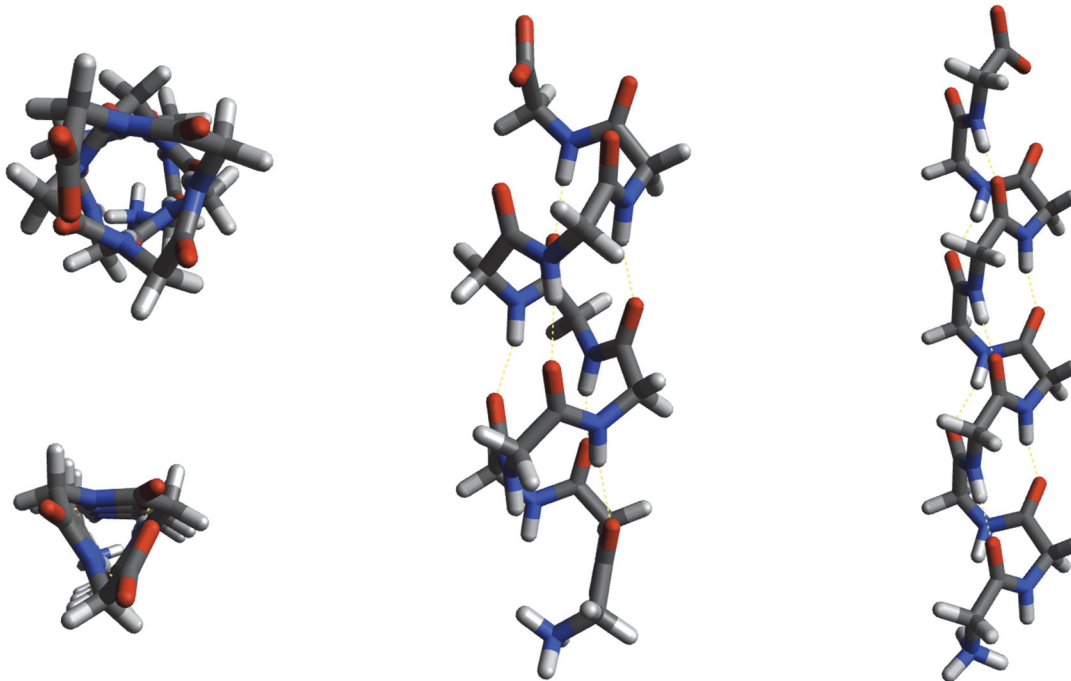
The  $\alpha$ -helix model is the result of powerful deductions about the chemistry of amide bonds and peptides by Pauling [1]. First, amide bonds were planar due to resonance, a fact not appreciated by others in the field (see Eisenberg [2]); second, intramolecular hydrogen bonding needed to be optimized; and third, there was no reason to restrict a model to an integral number of residues per turn. In 1951, Pauling, Corey and Branson [1] applied these constraints in the context of regular protein sequences to yield the  $\alpha$ -helix, which originally had 3.7 residues-per-turn [3] and 13 atoms in a hydrogen-bonded ring (i.e.,  $\alpha$ -helix =  $3.7_{13}$ -helix). Two critical assumptions were: the concept of linear groups (i.e., a series of residues all having identical conformations), and the use of single, linear hydrogen bonds (Fig 1) to fold the sequence in space by imposing a distance constraint between two non-consecutive amino acids in space. In one of the subsequent nine papers Pauling and Corey published in 1951, they suggested, “The number of residues per turn is fixed primarily by the bond angle at the  $\alpha$  carbon atom; it varies from 3.60 for bond angle  $108.9^\circ$  to 3.67 for bond angle  $110.8^\circ$ ”. The backbone torsional angles ( $\Phi = -57^\circ$ ,  $\Psi = -47^\circ$ ) associated with the  $\alpha$ -helix ( $3.6_{13}$ -helix) became official in the article by the IUPAC-IUB Commission on Biochemical Nomenclature published in 1970 [4]. These values were based on experimental fitting of parameters to diffraction data for poly-L-alanine in 1967 by Arnott and Dover [5].

In 1953, Donohue applied similar constraints with slightly different parameters to arrive at an alternative helical configuration (Fig 2), one with a rise of 3 residues-per-turn and 10 atoms ( $3_{10}$ -helix) in the hydrogen-bonded ring [7]. Where the  $\alpha$ -helix places residues  $i$  and  $i+4$  with a hydrogen bond, the  $3_{10}$ -helix is a longer and more tightly wound configuration with hydrogen bonds between residues  $i$  and  $i+3$  (Fig 2). Némethy et al. analyzed the equations derived by Miyazawa [8] regarding helical parameters (helix pitch ( $p$ ) and residues per turn ( $n$ )) and proposed the  $\alpha_{II}$ -helix in 1967 that exactly fit the helical parameters of the Pauling  $\alpha$ -helix, but lacked amide hydrogen bonds [9]. As the importance of amide H-bonds and three-centered



**Fig 1. Hydrogen-bonding alternatives; linear H-bond assumed by the Pauling/Donohue groups; three-centered hydrogen bonds to carbonyl oxygens that dominate crystal structures of small organic molecules [6].**

doi:10.1371/journal.pone.0123146.g001



**Fig 2. Schematic stick-figure diagrams of  $\alpha$ -helix (top left and middle) and  $3_{10}$ -helix (bottom left and right) helices.**

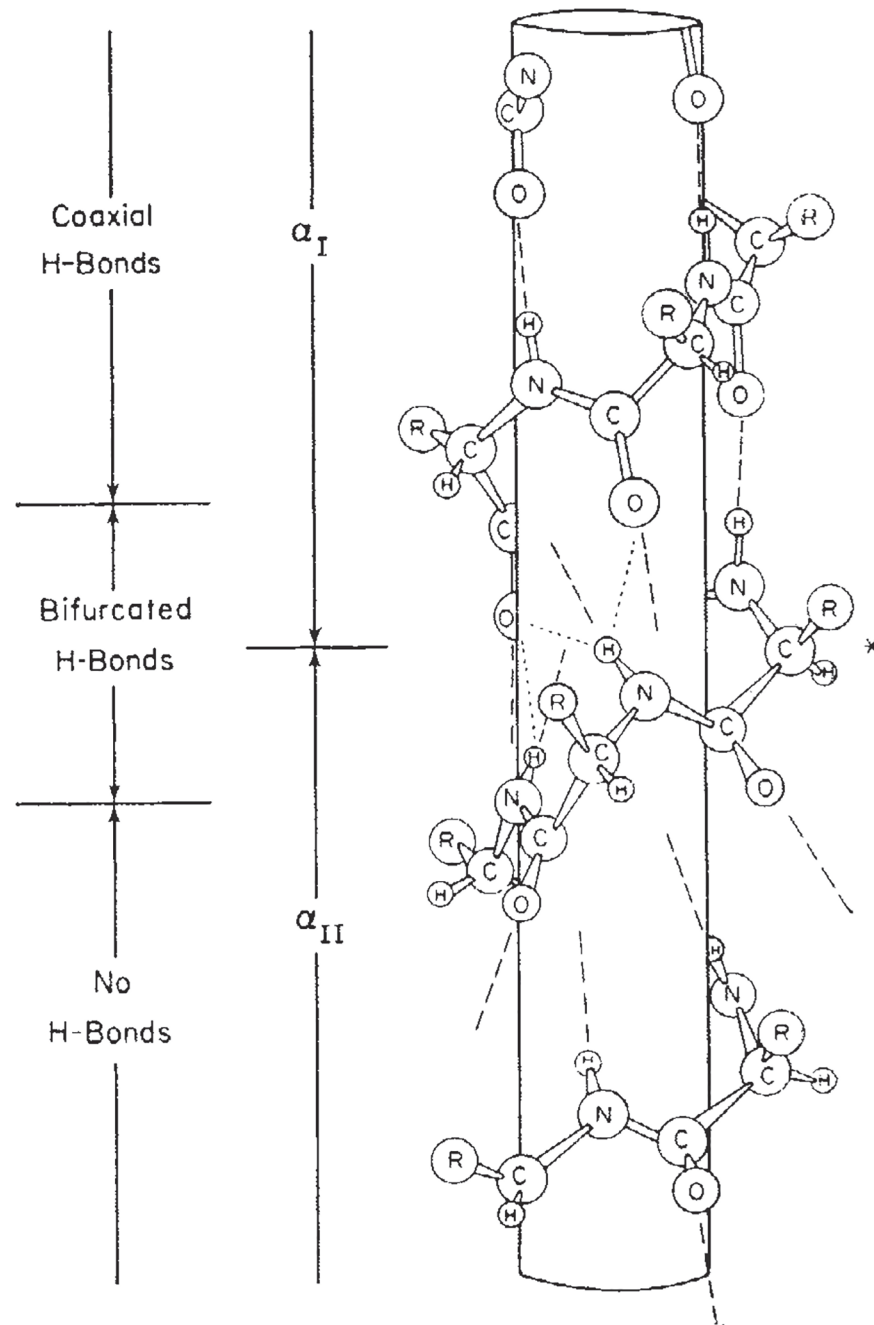
doi:10.1371/journal.pone.0123146.g002

hydrogen bonding (Fig 1) was just beginning to be appreciated [10], the  $\alpha_{II}$ -helix disappeared from the literature. In the Némethy et al paper [9], however, an intermediate helix between the  $\alpha$ -helix and the  $\alpha_{II}$ -helix was illustrated (Fig 3) with bifurcated hydrogen bonds and the same helical pitch and residues-per-turn as the Pauling  $\alpha$ -helix. These are exactly the properties observed for helices in high-resolution crystal structures from the Protein Data Bank (PDB) [11] that we refer to as Némethy-, N- or  $3.6_{13/10}$ -helices in this paper. The crucial insight by Némethy et al. was recognition that helical parameters are not tightly coupled to backbone torsional angles.

The canonical  $\alpha$ -helix and  $3_{10}$ -helix, which are the result of logically applying simple but powerful assumptions, laid the foundation for structural biology and dominated the interpretation of experimental data for over half a century. Whereas Pauling, Donahue and Némethy had access to very little structural data, today an abundance of structural data is available [12]. Armed with improved experimental data from over 100,000 protein structures, it is possible to revisit assumptions and refine parameters for protein helices without imposing restrictive assumptions.

### Classical assumptions vs. high-resolution data

In the intervening half-century, detailed understanding of molecular interactions as well as the surfeit of precise experimental data has led to concern about the early assumptions that linear hydrogen bonds are optimal electrostatic interactions between amides, and that linear groups apply to proteins. Moreover, early crystallographic data were often limited in resolution and employed external force field and/or secondary structural constraints during refinement; in turn, force fields were validated on crystallographic data, a “circular reference” potentially leading to propagation of errors. More sophisticated methods developed for small-molecule structures exemplified by REFMAC [13] or SHELXL [14] only began to take root for structural



**Fig 3. Reproduction of Fig 1 from Némethy et al. [9] showing an intermediate N-helix ( $3.6_{13/10}$ ) with the helical parameters of the classical  $\alpha$ -helix and bifurcated amide H-bonds (used with permission).**

doi:10.1371/journal.pone.0123146.g003

refinement of proteins in the middle of the 1990s. MD simulations with a requisite force field (monopole electrostatics that reinforce linear hydrogen bonding) were introduced in 1991 by Axel Brünger and used widely [15]. It is, in fact, significant improvements in crystallographic technology and refinement in the last two decades that have provided the high-resolution protein structures on which this research is based. Nevertheless, REFMAC5 provides low-resolution refinement tools that utilize secondary-structural constraints when the experimental



observations are of poor quality [13,16]. Systematic changes in helical torsional angles from PDB structures when binned by crystallographic resolution imply systematic error in data derived from low-resolution electron diffraction data.

## Linear hydrogen bonds

Both Pauling et al. [1] and Donahue [7] assumed a linear hydrogen bond between the carbonyl oxygen and the amide hydrogen of amide bonds, in order to place geometrical constraints on their helical models (Fig 2). Pauling et al. “assumed that each nitrogen atom forms a hydrogen bond with an oxygen atom of another residue, with the nitrogen-oxygen distance equal to 2.72 Å, and that the vector from the nitrogen atom to the hydrogen-bonded oxygen atom lies not more than 30° from the N-H direction.” The 3.7<sub>13</sub>-helix described had the angle between the N-H vector and N-O vector of ~10° [1]. This assumption has been computationally perpetuated by the use of monopole electrostatics in first-generation molecular mechanics. Although linear hydrogen bonds were a reasonable first-order approximation in 1951, more accurate treatment of electrostatics supports three-centered hydrogen bonds as seen in high-resolution crystal structures.

It has been known for some time from crystal structures of small molecules, that carbonyl oxygens do not form linear hydrogen bonds [6] as the total electron density due to the lone pairs is not geometrically positioned [17] to stabilize such interactions. For example, the hydrogen-bonded interactions seen experimentally in water clusters cannot be adequately reproduced by monopole electrostatics; in that case, the angle is approximately 57°, far from linear [18]. In fact, accurate electrostatic potentials surrounding small molecules (such as water) cannot be faithfully represented by monopole electrostatics; the addition of dipole and quadrupole terms is required [19–22]. The force field models (CHARMM, AMBER, OPLS-AA, etc.) commonly used in molecular modeling of proteins [23] utilize monopole electrostatics that produce linear hydrogen bonds between amide hydrogens and carbonyl oxygens. This can be demonstrated by minimizing the formula (Jeans’ formula) expressing the potential energy  $V$  of interaction between two dipoles as shown in Eq (1). Taking the derivative and setting  $\cos\theta_{12}$  to zero gives the minimum of the angles between the dipoles at 0° or 180° (i.e., a linear hydrogen bond).

$$V(r) = -\mu_1 \cdot \mu_2 \cdot (\cos\theta_{12} - 3\cos\theta_1 \cdot \cos\theta_2) \quad 1$$

where  $\mu$  = dipoles vectors,  $r$  = distance between the dipoles,  $\theta$  = angle between the dipoles.

Hydrogen bonds are dominated by electrostatics, consequently, energy minimization or molecular dynamics (MD) simulations using force fields with monopole electrostatics will overemphasize helices with linear hydrogen bonds. Note that Pauling and Donahue did not choose linear hydrogen bonds (and thereby the canonical  $\alpha$ - and 3<sub>10</sub>-helical configurations) with the belief that hydrogen bonds must always be linear in helices, but rather because Coulombic electrostatics could be solved and interpreted analytically without computers, and because these simple potentials were adequate to explain the very limited amount of experimental data available at the time. Others have questioned probable helical bias in force fields commonly used by molecular mechanics [23,24]. As these first-generation force fields have been often used in analysis of experimental X-ray diffraction and NMR structures [15], there is valid concern regarding helical bias in the primary source of protein structural data, the PDB [11]. Similar observations regarding the discrepancies between theoretical/experimental observations and the results of molecular modeling of side-chain hydrogen bonds were made by Morozov et al. [25].

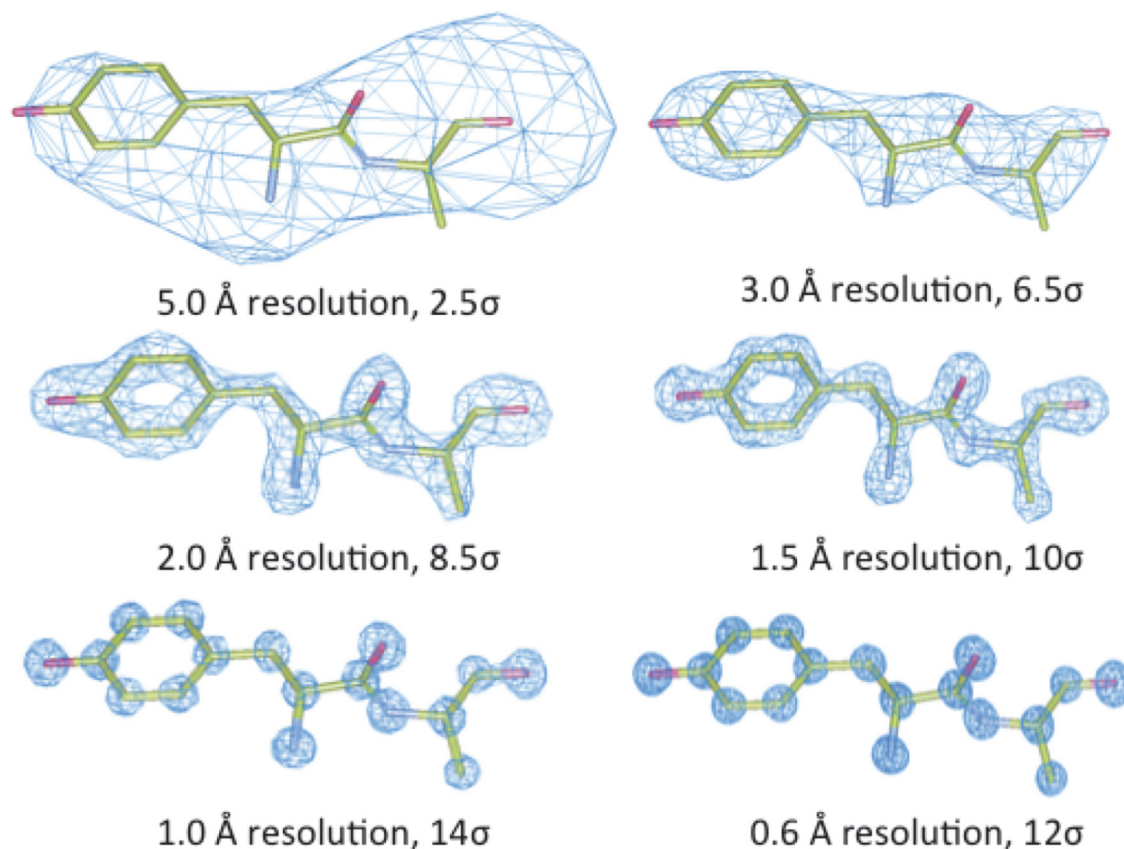
Recently, the limitation of monopole electrostatics has led to the development of second-generation force fields (such as AMOEBA [15,26]), which include both multipole electrostatics and polarizability for proteins and nucleic acids. The results obtained with AMOEBA are under scrutiny to validate the ability of improved electrostatics to predict experimentally measured thermodynamic quantities [27]. The relative orientation of water molecules in an AMOEBA simulation of water clusters matches that calculated by high-level *ab initio* quantum mechanical models, in contrast to the linear hydrogen bonding seen with monopole force fields [18]. AMOEBA has been evaluated for its ability to reproduce experimental electron density in peptide crystals [28] with only minor adjustments needed to get quantitative fits. Fenn et al. have shown recently that AMOEBA can correctly orient water networks and catalytically relevant hydrogens in proteins [29]. In a recent comparison of the ability of MD simulations to predict experimentally observed NOEs (nuclear Overhauser effects), AMOEBA was vastly superior to the three monopole force fields commonly used [30]. Thus, the computational tools to employ accurate physical approximations for electrostatics while exploring the effects of structural motifs on the dynamic behavior of proteins are now available [31].

### Linear groups assume that all amino acid residues in a helix have identical conformations

Another simplifying assumption employed by Pauling and Donahue was the presence of “linear groups”, defined as sequences of three consecutive amino acid residues all having the same  $[\phi, \psi]$ -angles  $\pm 10^\circ$ . The linear group assumption has not been supported by structural studies, and where they have been validated, the definition of “linear group” has been permissive enough to accommodate ensembles of similar-but-not-identical configurations. For example, according to a 2009 survey of high-resolution crystal structures of proteins by Hollingsworth et al. [32], only three distinct linear groups were found in 135 proteins. The most populated was a right-handed helical conformation dominated by the  $\alpha$ -helix, but also included the  $3_{10}$ -helix and mixed  $\alpha/3_{10}$ -helices. The other two linear groups detected were: a diverse group of extended conformations that are dominated by residues occurring in  $\beta$ -strands (in both parallel and antiparallel  $\beta$ -sheets), and a set of left-handed spiral conformations in the PII area.

### Implications of crystallographic resolution

Fortunately, the quality of protein crystal structures has improved dramatically over the past several decades. Early protein crystals diffracted poorly, and data collection was generally limited to a 2.5 Å resolution, or greater. Wlodower et al. illustrated the impact of nominal resolution on the electron density map from which structural models are derived [33] (Fig 4 shows a similar comparison based on crambin 3NIR). Whereas the highest resolution of 0.65 Å in the lysozyme structure employed data from 184,676 reflections for the electron-density map calculation, only 415 reflections were included at 5 Å resolution. Note that while higher-resolution maps provide enough geometric detail to unambiguously determine the amide plane and C = O vectors, there are multiple ways one could build a molecular model into density maps of poorer resolution [33]. X-ray crystallography is capable of localizing heavy atoms with astounding accuracy, so how is it possible for errors in helical populations to go undetected in so many crystallographic experiments? Interestingly, there is a straightforward two-part explanation: (i) hydrogens are less visible to X-ray diffraction due to their very small electronic cross-section; and (ii) the crystallographically determined positions of the heavy atoms, C $\alpha$ -C-N-C $\alpha$ , of the peptide backbone specify a unique helical configuration, i.e. helical parameters, the number of residues per turn and number of atoms in a hydrogen-bonding ring that was originally determined by fiber diffraction. The backbone torsional parameters ( $\Phi$ ,  $\psi$ ) for helical models



**Fig 4.** The appearance of electron density as a function of the nominal resolution of the experimental crystallographic data, (adapted from similar figure for the N-terminal fragment (Lys1—Val2—Phe3) of triclinic lysozyme (PDB: 2vb1) from Wlodawer et al. [33] for which permission to reproduce by the journal was not granted).

doi:10.1371/journal.pone.0123146.g004

having the same helical parameters, however, are determined by the orientation of the carbonyl  $C=O$  vector which have limited, but significant, degrees of freedom for the same helical parameters based on the insight of Némethy et al. [9].

The  $3_{10}$ -helix is elongated compared to the  $\alpha$ -helix (Fig 2), a consequence of its tighter helical twist that is imposed by  $i \rightarrow i+3$  vs.  $i \rightarrow i+4$  hydrogen bonding pattern for the  $\alpha$ -helix. As the  $3_{10}$ -helix is longer than the  $\alpha$ -helix, one might have expected changes in helical parameters (i.e., different number of residues-per-turn, helical pitch, position of  $C\alpha$ s and orientation of  $C\alpha$ - $C\beta$  vectors) as the backbone torsional angles moved between those of the classical  $\alpha$ -helix towards that of the  $3_{10}$ -helix. However, the torsional parameters for  $\phi$  and  $\psi$  can change significantly without perturbing  $C\alpha$  atomic positions as pointed out by Némethy et al. in 1967 [9]. Crankshaft motions of torsional angles reorient the peptide plane and amide carbonyl without changing positions of  $C\alpha$ - $C\beta$  atoms, i.e. the helical parameters. Whereas small changes in backbone torsion angles change the orientation of the amide bond to allow amide three-centered hydrogen bonds to optimize electrostatic interactions, unless the resolution is sufficient to resolve the position of the carbonyl oxygen within the helical cylinder, there is little difference during refinement of a molecular model to fit experimental electron density.

Determining an accurate position and orientation of the carbonyl oxygen of the peptide bonds is difficult (Fig 4) until the resolution is less than approximately 2.0 Å. Because hydrogens are less-well resolved by X-ray scattering in protein crystallography, the H-N amide

orientation is effectively invisible, and backbone torsional  $[\phi, \psi]$  parameters depend on the detailed orientation of the modeled carbonyl C = O vectors. For data gathered at poorer resolutions, fine-grained details were irrelevant, and the classic  $\alpha$ -helical cylinder, refined with the assistance of monopole force fields, was assumed to fit the ambiguous geometric constraints. Thus, models built from low-resolution data ( $>2.0 \text{ \AA}$ ) are susceptible to introduction of bias from modeling assumptions, and (in the absence of experimental constraints) might be expected to model classic linear groups. Models built from high-resolution data ( $<2.0 \text{ \AA}$ ) are sufficiently constrained by experimental factors to model the detailed configurations of helical backbones without imposing such limiting assumptions.

### High-resolution models have less bias than poorer-resolution models

Careful analyses of high-resolution ( $<2.0 \text{ \AA}$ ) protein crystallographic models have, for some time, revealed discrepancies between models supported by accurate experimental data and the classical helical models. More than thirty years ago, Blundell et al. [34,35] recognized how the solvent-shielded interior of a helix tended to have longer and less linear hydrogen bonds, “. . . recent high-resolution (1.7 – 2.0  $\text{\AA}$ ) diffraction studies have shown that the parameters described by Pauling et al and later by Perutz [35] and by Arnott and Wonacott [36] are not a good description of the  $\alpha$ -helices in globular proteins, where the mean values of  $\phi, \psi$  are usually close to  $-63$  degrees,  $-42$  degrees. . . and arose as a mean of two significantly different classes in amphipathic helices depending on whether the peptide carbonyl oxygen is hydrogen bonded to a solvent or polar side-chain atom. The hydrogen bonds made by the hydrophilic carbonyls to the NH groups within helices are longer and less linear than those involving hydrophobic carbonyls.” [34]

## Results and Discussion

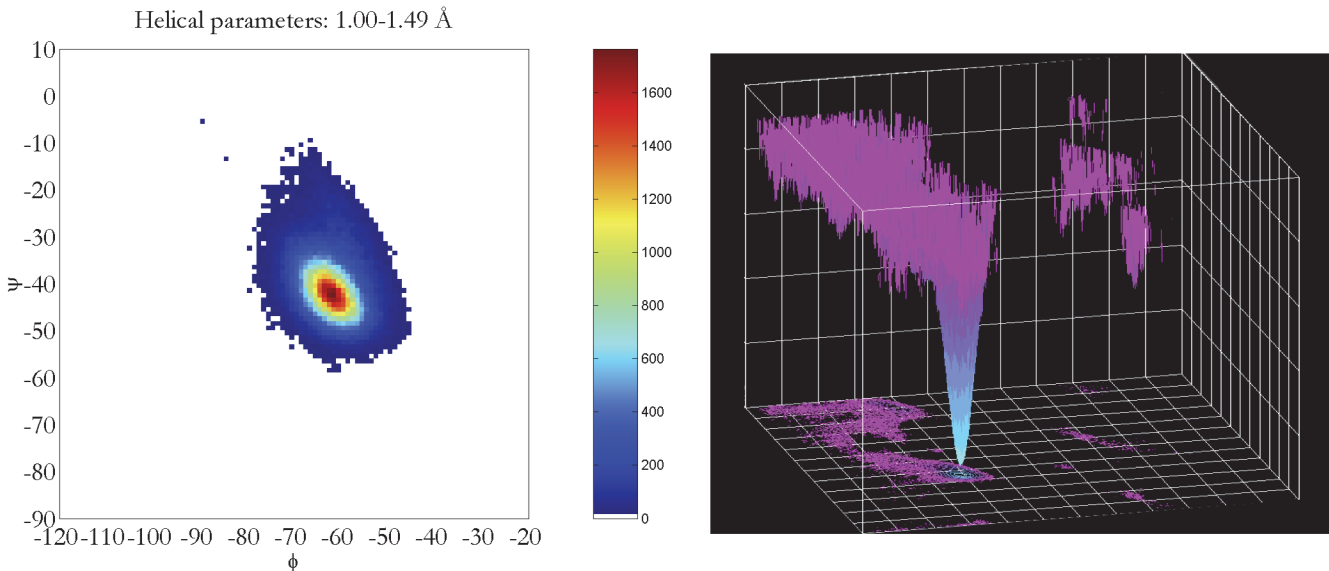
### Analysis of protein helices binned by nominal crystallographic resolution

Analysis of all experimental structural models of helices in proteins in the Protein Databank (Table 1) exposed subtle inconsistencies in the torsional angle ( $\Psi, \phi$ ) parameters, when comparing high-resolution data against the classic  $\alpha/3_{10}$ -helical models. At high resolution, where the experimental electron density accurately constrained helical conformations without the need for any structural assumptions or force-field constraints [33], the backbone torsional angles populate a smooth distribution of helical conformations with a single maximum (Fig 5) at  $\phi = -62^\circ, \Psi = -43^\circ$  backbone torsion angles rather than the anticipated bimodal distribution of  $\alpha$ - and  $3_{10}$ -helices [37]. Hovmoller et al. had previously analyzed a non-redundant set of 1042

**Table 1. Hydrogen-bond analysis of the PDB as a function of nominal crystallographic resolution.**

| Resolution | Total H-Bonds | $\alpha$ -Helix Bonds | $3_{10}$ -Helix Bonds | $\beta$ -Sheet Bonds | Shared 3-center H-bonds |
|------------|---------------|-----------------------|-----------------------|----------------------|-------------------------|
| 0.00–0.49  | 101188        | 34594                 | 14873                 | 20531                | 14283                   |
| 0.50–0.99  | 33446         | 9693                  | 2345                  | 12329                | 1980                    |
| 1.00–1.49  | 718464        | 246901                | 57088                 | 230275               | 47186                   |
| 1.50–1.99  | 5342073       | 1942587               | 451007                | 1595126              | 375252                  |
| 2.00–2.49  | 7739951       | 2936903               | 719304                | 2157150              | 626744                  |
| 2.50–2.99  | 5178485       | 1922736               | 537574                | 1458120              | 505260                  |
| 3.00–3.49  | 1839170       | 672369                | 202288                | 493697               | 210678                  |
| 3.50–3.99  | 406006        | 141464                | 44390                 | 109627               | 51326                   |
| 4.00–4.49  | 137350        | 51589                 | 15757                 | 32724                | 18208                   |
| 4.50–4.99  | 32332         | 12830                 | 4077                  | 6711                 | 4205                    |

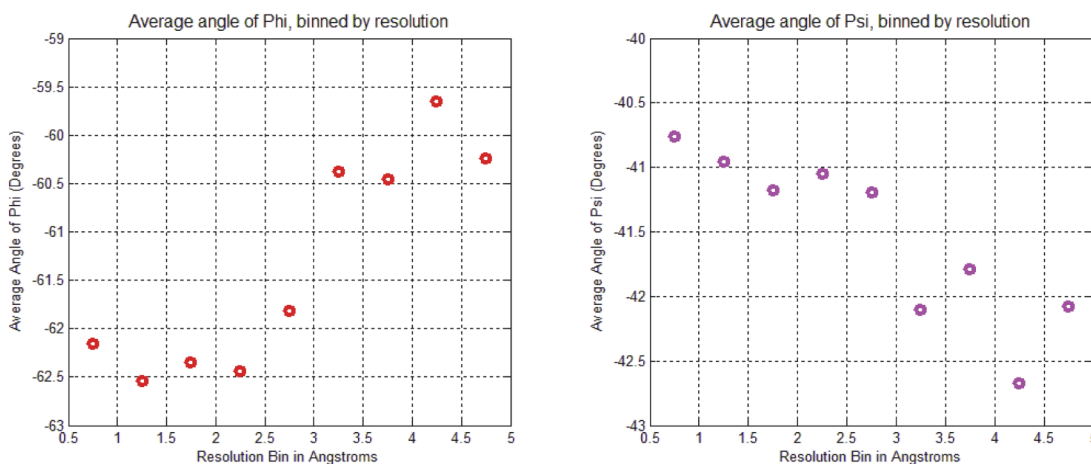
doi:10.1371/journal.pone.0123146.t001



**Fig 5. Left: Distribution of helical backbone torsion angles ( $\Phi$ ,  $\Psi$ ) for 1.0–1.5 Å nominal resolution of 3,462 protein structures from the PDB. Right: Conversion of density of high-resolution helices to a relative energy scale by Boltzmann-weighting the distribution.**

doi:10.1371/journal.pone.0123146.g005

protein subunits from the PDB (January 3, 2002) with regard to secondary structure. The proteins were all determined by X-ray diffraction to a resolution of 2.0 Å or higher, refined to  $R \leq 0.20$  with less than 30% sequence homology [38]. When five or more consecutive amino acids had torsion angles in the  $\alpha$ -helical region of the Ramachandran plot, the hydrogen bonds typical of the  $\alpha$ -helix were formed, and the resulting structure fit a very highly populated and finely focused area near  $\phi = -63.8^\circ$ ,  $\Psi = -41.1^\circ$ . No discussion of the implications of these non-classical torsional parameters on helical rise-per-residue or number of residues-per-turn appeared. A helix with  $\phi = -62^\circ$ ,  $\Psi = -43^\circ$  backbone torsion angles, bifurcated hydrogen bonds, and the helical parameters of the  $\alpha$ -helix (Fig 3) is herein referred to as the experimental Némethy-, N- or 3.6<sub>13/10</sub>-helix. By examining the backbone torsional angles by nominal resolution, a compensatory shift in  $\Phi$ ,  $\psi$  was observed (Fig 6). This purely experimental perspective



**Fig 6. Average values of  $\Phi$  (Phi) and  $\Psi$  (Psi) torsion angles of helical residues in PDB binned by nominal crystallographic resolution. This compensatory change in  $\Phi$ ,  $\Psi$  is due to a “crankshaft” motion [39] of the amide bond that maintains the relative positions of the  $\alpha$ -carbons of the peptide backbone despite significant deviations in  $\Phi$ ,  $\Psi$  torsions.**

doi:10.1371/journal.pone.0123146.g006



of protein helices determined at high resolution is consistent with bifurcated (three-centered) hydrogen bonds with a strong hydrogen bond to the  $i+4$  residue simultaneous with a weaker  $i+3$  hydrogen bond rather than the single linear hydrogen bond assumed by the classical helical models. What is surprising is that the observation of different average backbone torsional angles for protein helices was first reported by Blundell et al. in 1983 [34], and subsequently thereafter [38], without to our knowledge ever plotting the distribution of  $\phi$  and  $\Psi$  (Fig 5).

### Averaged bimodal population vs. a single continuous population

In contrast to the usual expectation that the observation of non-classical backbone torsional angles was due to averaging of a bimodal distribution of many  $\alpha$ -helices with a few  $3_{10}$ -helices, this study shows how the distribution of torsion angles in high-resolution torsion angles populated a single, continuous, smooth peak that is intermediate between the  $\alpha$ - and  $3_{10}$ -helical forms. Whereas a bimodal population would be the expected result for helices that have some proportion of  $i \rightarrow i+4$  hydrogen bonds (i.e.,  $\alpha$ -helical) and some proportion of  $i \rightarrow i+3$  hydrogen bonds (i.e.,  $3_{10}$ -helical), a single population is consistent with helices that are positioned to transition easily between  $i \rightarrow i+3$  and  $i \rightarrow i+4$  hydrogen bonds, and indeed have a significant portion with simultaneous three-center shared  $i \rightarrow i+3$  and  $i \rightarrow i+4$  hydrogen bonds. The evidence is consistent with a single consensus helical ensemble, the Némethy-, N- or  $3.6_{13/10}$ -helix, and a single local energy minimum, rather than two separate helical forms. This observation of non-classical helical torsion angles has appeared in the literature numerous times [38,40,41], but structural implications have apparently been largely ignored.

### Straight-forward explanations for persistence of an inaccurate helical models

The longevity and dominance of these two canonical helices in protein science is quite amazing, considering that two of Pauling's key assumptions have been experimentally shown to be inaccurate. Bifurcated three-centered hydrogen bonds in proteins were found to occur with a frequency of 24% in 13 high-resolution ( $<1.6 \text{ \AA}$ ) structures by Preissner et al. in 1991 [40]. McDonald and Thornton examined a dataset of 57 higher-resolution structures ( $<2.0 \text{ \AA}$ ) in 1994 and classified hydrogen bonding [42]. More significantly, the paper included diagrams and energetic estimates of three-centered hydrogen bonds by quantum mechanics with distributed multipole analyses. According to this analysis, four out of every five-backbone carbonyls failed to form three-centered hydrogen bonds, implying a frequency of 20% occurrence. This work was extended by Fain et al. in 2001/2008 on the presence of double and/or shared hydrogen bonds (estimated at approximately two-thirds of observed backbone hydrogen bonds and supported by quantum calculations) in the helices of globular proteins [43,44]. The discrepancy between the two studies probably arose from differences in the resolutions of the datasets examined and/or the choice of H-bonding parameters. Nevertheless, it is quite clear from these studies that amide three-centered hydrogen bonds are common features of high-resolution crystal structures of proteins and contradict the assumption of linear hydrogen bonds. Many have questioned the use of monopole electrostatics in molecular mechanics as misleading [20–22,45]. Use of multipole electrostatics and polarizability in X-ray refinement leads to significant improvements in the fits between model and electron density [28,29]. Cardamone et al. have recently published a critical review of multipole electrostatics, and their advantages over monopoles [46]. Nevertheless, protein crystal structures, from the originals by Kendrew [47] and Perutz [48] to recent additions to the PDB, have continued to utilize the two classic helices in their interpretation of electron density.



## Materials and Methods

### Hydrogen-bond statistics of the Protein Data Bank

Files were downloaded from the Protein Data Bank (PDB). A local PDB archive was created in 2008 and updated continuously through June 9, 2010, when this analysis was performed. A total of 60,755 proteins were examined, of which 53,040 were determined by X-ray crystallography (the main focus of this study); the remaining structures were determined by nuclear magnetic resonance (NMR) methods, electron microscopy, homology modeling, and other hybrid methods. Each structure was binned by nominal crystallographic resolution, cleaned, and pre-processed to fix structural anomalies and format inconsistencies. The 3D coordinates of main chain heavy atoms (i.e., 'ATOM' records) were not altered. Each of the resolution bins were then analyzed using programs written or adapted by Eric Welsh and Daniel J. Kuster as part of the StructureTools package [49], available from the authors under an open source (GPL) license. These scripts analyzed each crystallographic PDB file and generated a detailed list of hydrogen bonds. Secondary structures were then assigned based on the hydrogen-bonding patterns for each protein. Every helical hydrogen bond within each resolution bin was assigned to a category to give a total statistic for each resolution bin of the PDB as seen in Tables 1 and 2.

### Hydrogen-bond detection

Hydrogens were added to each structure using the HBPLUS program of McDonald and Thornton [42], which was chosen after quantitative comparison of hydrogen bond geometries and visual inspection of several thousand hydrogen-bond placements for various algorithms. These criteria for hydrogen bonds would be relatively permissive for hydrogen bonds at large, but were well suited for the solvent-shielded interior of a protein helix. No extra constraints were placed on the detection of multiple hydrogen bonds—if atoms were close enough and within planar angles where hydrogen bond could occur, a hydrogen bond was detected. As a quality check, thousands of hydrogen-bond pairs were visually inspected using the SYBYL molecular modeling program; the severe steric constraints of helical backbone chain atoms made detection of geometric anomalies a straightforward (if tedious) task. Both random samples and a ranked list of the most extreme geometries were visually inspected to validate the procedure.

### Hydrogen-bond patterns define helices

Helices were assigned based on the pattern of  $i \rightarrow i+3$  and/or  $i \rightarrow i+4$  hydrogen bonds; a helix must have at least two consecutive helical hydrogen bonds (i.e., the shortest possible helical hydrogen bond pattern would be  $i \rightarrow i+3$  followed by  $i+1 \rightarrow i+4$ ). The final geometric parameters were sufficient to detect hydrogen bonds with zero anomalies/false positives. Tables 1 & 2 shows the statistics for each category by nominal X-ray resolution.

### Helical backbone statistics in the PDB appear to be converged

Repeated analyses of the helical hydrogen bond content of the PDB from 2008 to 2010 resulted in essentially identical plots [37]. The fact that neither the shape of distributions, nor the statistical moments of the distributions changed over this time suggested that the population of helical backbones in the PDB (more than 10 million observations of helical hydrogen bonds), was statistically converged, and unlikely to change either the details of analysis or conclusions discussed herein by analysis of more structures.

**Table 2. Structural Percentage of the PDB as a function of nominal crystallographic resolution.**

| Resolution | $\alpha$ -Helix Bonds % | $3_{10}$ -Helix Bonds % | $\beta$ - Sheet Bonds % | Shared Bonds % |
|------------|-------------------------|-------------------------|-------------------------|----------------|
| 0.00–0.49  | 34.19%                  | 14.70%                  | 20.29%                  | 14.11%         |
| 0.50–0.99  | 28.98%                  | 7.01%                   | 36.86%                  | 5.92%          |
| 1.00–1.49  | 34.37%                  | 7.95%                   | 32.05%                  | 6.57%          |
| 1.50–1.99  | 36.36%                  | 8.44%                   | 29.86%                  | 7.02%          |
| 2.00–2.49  | 37.94%                  | 9.29%                   | 27.87%                  | 8.10%          |
| 2.50–2.99  | 37.13%                  | 10.38%                  | 28.16%                  | 9.76%          |
| 3.00–3.49  | 36.56%                  | 11.00%                  | 26.84%                  | 11.46%         |
| 3.50–3.99  | 34.84%                  | 10.93%                  | 27.00%                  | 12.64%         |
| 4.00–4.49  | 37.56%                  | 11.47%                  | 23.83%                  | 13.26%         |
| 4.50–4.99  | 39.68%                  | 12.61%                  | 20.76%                  | 13.01%         |

doi:10.1371/journal.pone.0123146.t002

### Ramachandran plot of helical configurations by resolution

The classic two-dimensional Ramachandran-style plots of  $[\phi, \psi]$  torsion angles were produced using a Matlab script. Millions of helical hydrogen-bond observations were clustered into  $1^\circ \times 1^\circ$  bins, allowing fine-grain frequency statistics to be calculated and visualized across all  $[\phi, \psi]$  space. Distributions of  $\omega$  torsion angles were also examined. As expected due to the planar nature of the amide bond,  $\omega$  angles deviated very little from the  $180^\circ$  *trans*-conformation, so one can safely assume this invariant dimension does not contribute to helical backbone observations.

### Comparison with experimental neutron-diffraction data on proteins

Neutron diffraction allows the position of hydrogens to be determined more precisely. In contrast to X-ray diffraction experiments where the position of electron-sparse hydrogens are inferred from the coordinates of heavy atoms, neutron diffraction images hydrogen nuclei directly. The PDB had over 20 proteins structures determined by neutron diffraction. Improvements to hydrogen-bond orientation/geometry were seen when two DNA and a protein crystal structure were refined against joint neutron/X-ray diffraction data sets using a force field derived with AMOEBA multipole electrostatics [50].

Different proteins (PDB Codes = 1L2K, 1L2N, 1NTP, 1VCX, 1WQ2, 2DXM, 2GVE, 2INQ, 2MB5, 2R24, 2VS2, 2WYX, 2YZ4, 2ZOI, 2ZWB, 2ZYE, 3A1R, 3BYC, 3HGN, 3INS, 3KKX, 3KMF, 3KYX, 3L45, 5PTI and 5RSA) whose structures had been determined by neutron diffraction were analyzed for their hydrogen-bonding patterns. This set contained 3659 backbone H-bonds in the 26 proteins. Structure Tools [49] was used to detect and analyze hydrogen-bond content according to geometric rules. As can be seen by the statistics in Table 3, shared H-bonds account for 18.5% of the overall total of 1711 helical H-bonds detected. Thus, although the studies by neutron diffraction studies are limited compared to X-ray crystallography, the ability of neutron diffraction to directly locate the amide hydrogen strongly supported a significant role for shared H-bonds in protein helices.

### Methods for Calculating Helical Parameters

The mathematics used to analyze the impact of changes in backbone torsion angles on helix pitch and orientation of  $C\alpha$ - $C\beta$  vectors in proteins was described by Ramachandran et al. [51]. A more recent and thorough generalization has been made based on quaternions [52]. The

**Table 3. Hydrogen-bond analyses for neutron-diffraction data for 26 proteins.**

| Total H-Bond Summary                 | Total H-Bond Summary By Percentage                 |
|--------------------------------------|--|
| Total i to i+4 bonds: 1329           | Total i to i+4 bonds: 36.3%                        |
| Total i to i+3 bonds: 382            | Total i to i+3 bonds: 10.4%                        |
| Total H-Bonds in Helices = 1711      |  |
| Total β-Sheet bonds: 1043            | Total β-Sheet bonds: 28.5%                         |
| Total # of bonds at bends: 431       | Total # of bonds at bends: 11.8%                   |
| Total # of bonds at turns: 26        | Total # of bonds at turns: 0.7%                    |
| Total # of bonds at coils: 302       | Total # of bonds at coils: 8.3%                    |
| Total Disallowed bonds: 146          | Total Disallowed bonds: 4.0%                       |
| Summary—3659 H-bonds detected = 100% |  |
| Bifurcated H-bonds: 317              | % Bifurcated H-Bonds in Helices = 317/1711 = 18.5% |

doi:10.1371/journal.pone.0123146.t003

following equations were used to analyze the impact of the backbone torsion parameters  $\Phi$ ,  $\Psi$  on helix pitch ( $p$ ) and residues per turn ( $n$ ) (Figs 7, 8, 9 and 10):

Model equations used for helix-winding relationships:

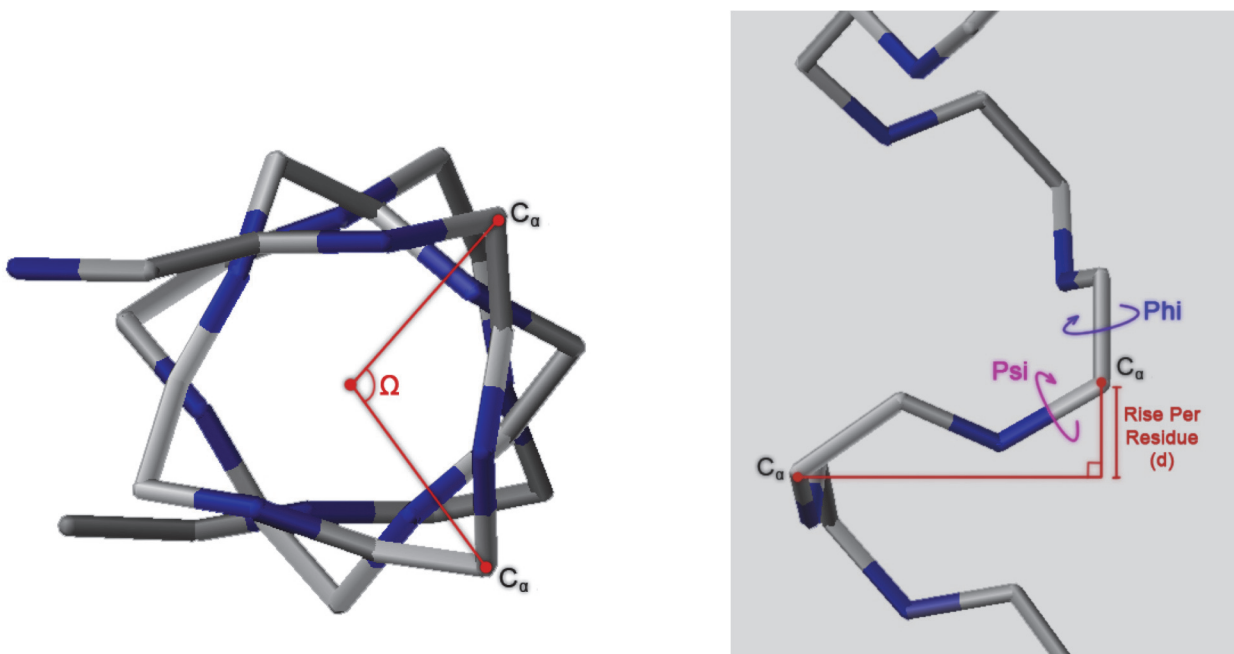
Helical pitch ( $p$ , rise per turn) =  $d$  (Rise per residue) \*  $n$  (Residues per Turn)

Relationships exists between  $\Phi, \psi$  and  $\Omega$  for calculating  $n$  and  $d$ .<sup>1</sup>

$$s = \frac{\Phi + \Psi}{2}, t = \frac{\Phi - \Psi}{2}$$

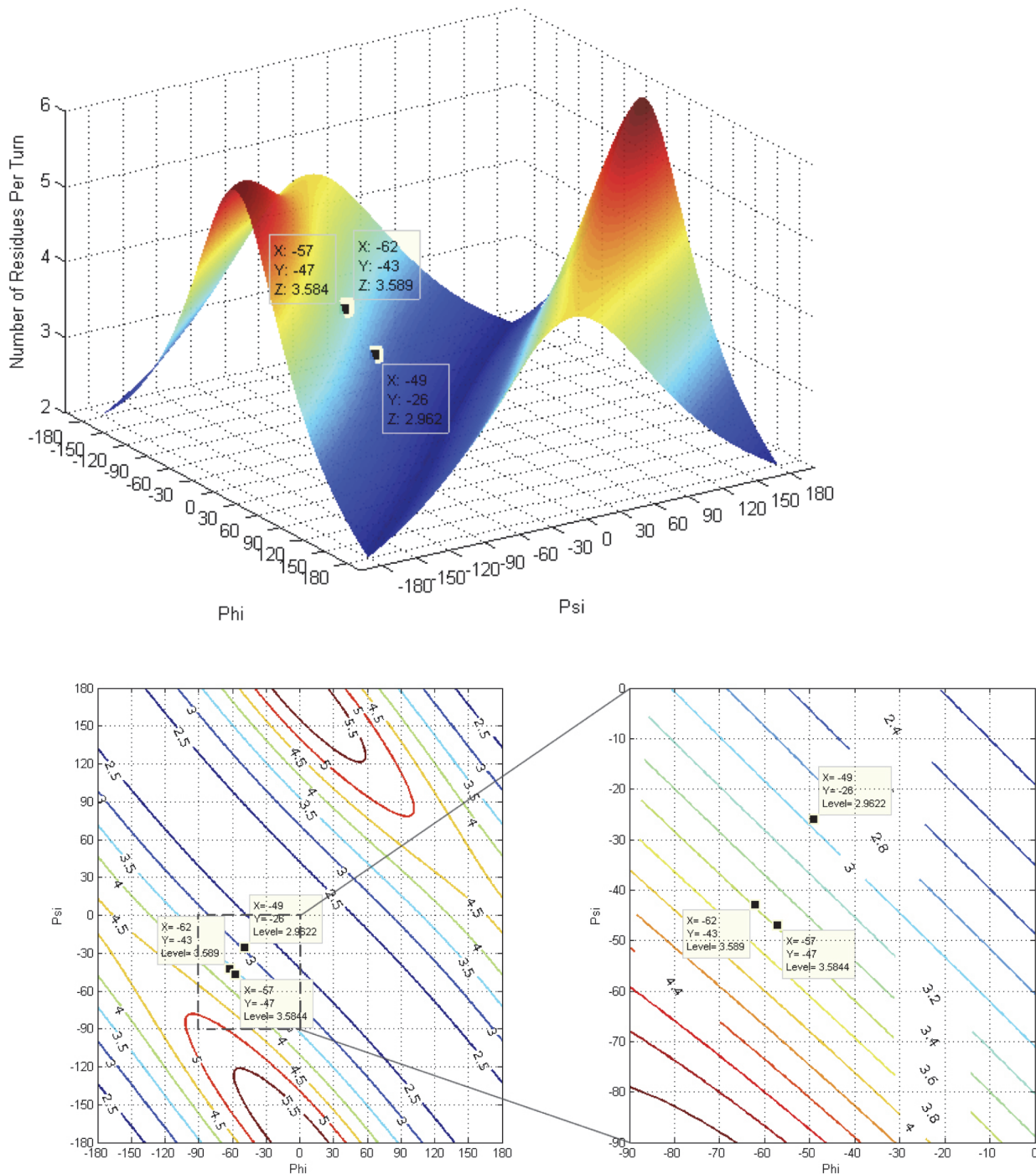
$$\cos\left(\frac{\theta}{2}\right) = -0.817\sin(s) + 0.045 \sin(t)$$

2



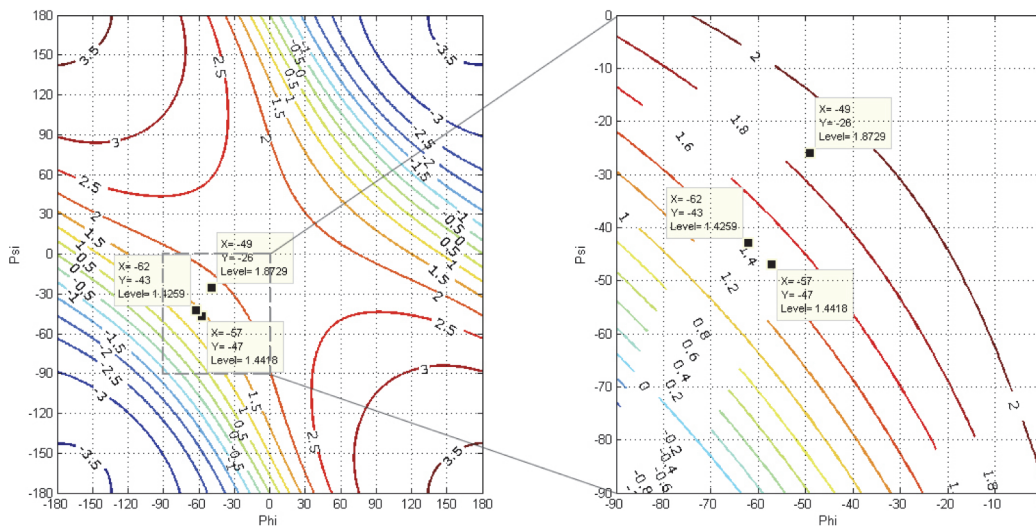
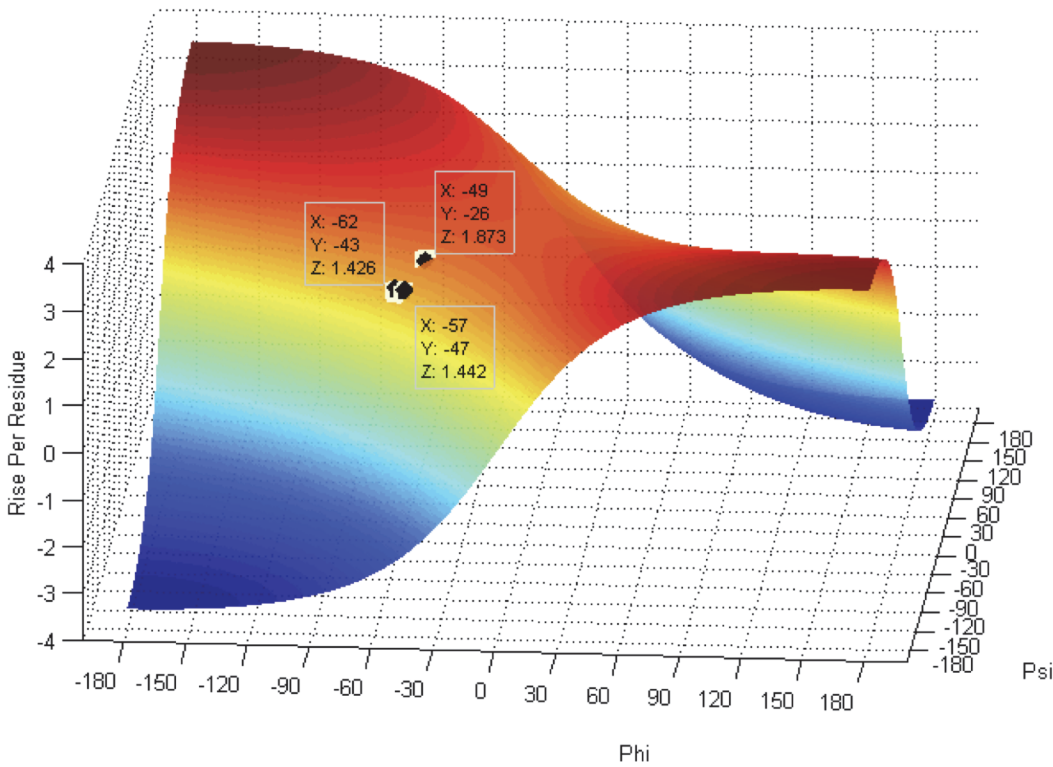
**Fig 7. Definition of helical parameters with an oligopeptide helix backbone (nitrogens in blue).** Left: Helix winding ( $\Omega$ ) is the angle between two adjacent  $C\alpha$  vectors ( $n = \text{residues-per-turn}/360$ ). Right: Helix pitch ( $p$ ) is the rise-per-residue, or vertical distance between  $\Omega$  adjacent residues, projected on the helical axis.

doi:10.1371/journal.pone.0123146.g007



**Fig 8. Contours and 3D surface helical pitch ( $n$ ) overlaid on 2D  $\phi$ ,  $\psi$  plot.** Positions of classic  $\alpha$ -helix ( $\phi = -57$ ,  $\psi = -47$ ),  $3_{10}$ -helix ( $\phi = -49$ ,  $\psi = -26$ ), and experimental Némethy-, N- or  $3.6_{13/10}$ -helix ( $\phi = -62$ ,  $\psi = -43$ ) are indicated. Note that the helical pitch is nearly identical for the  $\alpha$ - and  $3.6_{13/10}$ -helices—they lie on a contour of essentially equal value.

doi:10.1371/journal.pone.0123146.g008



**Fig 9. Contours and 3D surface showing rise-per-residue ( $d$  in Å) overlaid on 2D  $\phi, \psi$  plot.** Positions of classic  $\alpha$ -helix ( $\phi = -57, \psi = -47$ ),  $3_{10}$ -helix ( $\phi = -49, \psi = -26$ ), and Némethy-, N- or  $3.6_{13/10}$ -helix ( $\phi = -62, \psi = -43$ ) are indicated. Note that  $d$  is approximately the same for the  $\alpha$ - and  $3.6_{13/10}$ -helices.

doi:10.1371/journal.pone.0123146.g009

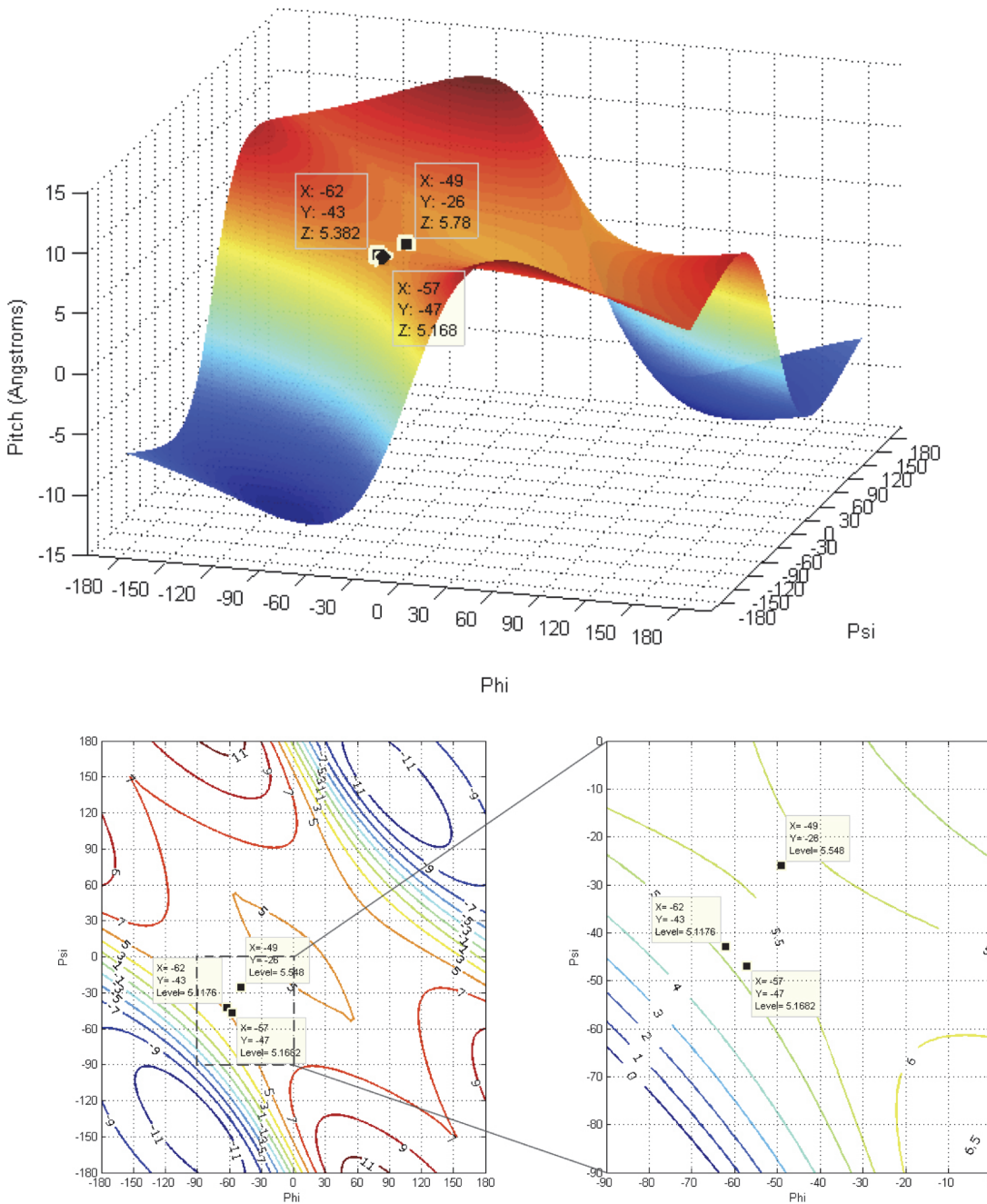
$\theta$  can then be used to calculate  $n$  and  $d$ :

$$n = 360/\theta$$

$$d \sin\left(\frac{\theta}{2}\right) = -0.68\cos(t) + 2.9 \cos(s)$$

3





**Fig 10. Contours and 3D surface showing number of residues-per-turn ( $n$ ) overlaid on 2D  $\phi$ ,  $\psi$  plot.** Positions of classic  $\alpha$ -helix ( $\phi = -57$ ,  $\psi = -47$ ),  $3_{10}$ -helix ( $\phi = -49$ ,  $\psi = -26$ ), and experimental Némethy-, N- or  $3.6_{13/10}$ -helix ( $\phi = -62$ ,  $\psi = -43$ ) are indicated. Note that  $n$  is approximately the same for the  $\alpha$ - and  $3.6_{13/10}$ -helices.

doi:10.1371/journal.pone.0123146.g010

- $\theta$  — Rotation Angle per Residue in Trans
- $\phi$  — Phi, Torsion Angle of N- $C_{\alpha}$
- $\psi$  — Psi, Torsion Angle of  $C_{\alpha}$ -C'



Thus, helical pitch ( $p$ ) can be expressed directly in terms of  $\Psi$  and  $\Phi$ .

$$p = d \times n$$

- <sup>1</sup>  $\theta$ ,  $\Psi$ , and  $\Phi$  are expressed in radians.
- <sup>2</sup> resulting pitch is expressed in Angstroms per turn.

## Force-field dependence

Because peptides are so inherently flexible and multiple helical forms have been suggested depending on the experimental/theoretical approach, the choice of force field for the MD simulations was crucial. In particular, force fields dependent on monopole electrostatics only reinforce Pauling's assumption of linear hydrogen bonds. A previous study by Zheng et al. [30] compared three monopole force fields in common use with AMOEBA, a second-generation force field using multipole electrostatics and polarizability [26]; a significant improvement in the ability of the AMOEBA force field to predict experimental observations on peptide dynamics was observed and justified the choice of AMOEBA for the MD simulations.

## Isothermal-isobaric MD simulations

All solvent-solute systems were optimized via the TINKER MINIMIZE program with an 0.01 kcal/mol/Ang RMS gradient cutoff prior to dynamics simulation. All MD simulations were performed using either the Ponder lab compute cluster or compute nodes at the Washington University Medical School Center for High Performance Computing. Since the replica-exchange MD simulation only offered structure samplings, rather than true dynamic trajectories, a total of 8 isothermal-isobaric simulations for oligoalanine helices at 300° K and under 1.0 atmospheres were run for 17 ns to obtain trajectory data. The vdw cutoff was set to be 12.0 Å and the particle-mesh Ewald real-space cutoff to 7.0 Å. Induced dipoles were converged to an RMS change of less than 0.01 Debye per iteration. The simple bonded/non-bonded RESPA integration scheme [53] was used to enable a larger simulation step size of 2.0 fs. Trajectory frames were recorded every 1 ps. The resulting hydrogen-bond matrix was compared to the matrix from the replica-exchange, and found to be reasonably similar suggesting that trajectory sampling was adequate.

## Replica-exchange MD

Replica-exchange MD simulation was performed with the protocol of Sugita and Okamoto [54] to get extensive structural sampling for the oligoalanine helices. Twenty-four replicas exponentially distributed from 300° K to 400° K were established with the help of an automated temperature generator for REMD-simulations [55] to obtain an exchange rate near 20% for each structure. The simulation was carried out with the TIREX replica-exchange program developed by Penev et al. [56]. All the setups were the same as the isobaric-isothermal simulations as described above, except the NVT ensemble was used and the trajectories were recorded every 0.5 ps. Each replica was simulated for 10 ns with exchanges attempted every 1000 steps.

## Molecular dynamics simulations of crambin in water [37]

The protein crambin from PDB entry 1CRN [57] was hydrated in a periodic box of explicit water. For the AMOEBA simulation, the AMOEBA water model [18] was employed. For

CHARMM and OPLS-AA simulations, the TIP3P water model [58] was used. Simulation data was collected over tens of nanoseconds and the resulting structures analyzed for changes in coordinates.

### AMOEBa molecular dynamics simulation of oligoalanine in water

The starting alanine helix of 12 residues was directly built using the TINKER PROTEIN program with  $\Phi = -60^\circ$  and  $\Psi = -43^\circ$  torsion angles to generate a right-handed  $\alpha$ -helix with an N-terminal ACE (acetyl) and a C-terminal NME (N-methylamide) capping groups. After minimization by TINKER MINIMIZE with the AMOEBa force field with a 0.01 cutoff, the final torsional angles were approximately ( $-63^\circ$ ,  $-39^\circ$ ) similar to ones observed in high-resolution crystal structures. The TINKER XYZEDIT program was used to tailor a water box to solvate the starting helix. Since the TIREX replica exchange program offered only NVT options, the box boundaries were manually adjusted to achieve an average pressure close to 1 atm at  $300^\circ\text{K}$  resulting in a water box with  $40.7777\text{ \AA}$  on each side.

### AMOEBa molecular dynamics simulation of oligo- $\alpha$ -aminoisobutyric acid (Aib) in water

The AMOEBa force-field parameters for Aib were constructed with the described protocol [27] by fitting to the Gaussian09 quantum calculation of an Ac-Aib-NHMe dipeptide. A right-handed, 12-residue,  $\alpha$ -helical capped oligo- $\alpha$ -aminoisobutyric acid was built the same way as the oligoalanine helix. The water box was set to be  $40.787\text{ \AA}$  on each side to achieve a pressure close to 1 atm. The replica exchange, trajectory extraction, and Ramachandran map were done the same as for oligo-alanine mentioned above.

### Energy surface sampling in the $\alpha$ -helical region

Ac-Ala-Ala-Ala-NMe structures in the range of  $\phi = [-100$  to  $-50]$  and  $\psi = [-50$  to  $0]$  were constructed with TINKER PROTEIN package at 1 degree increments for force field and 2 degree increments for quantum mechanics calculations. Then, their energy was computed with AMOEBa09, OPLS-AA, OPLS-AAL, CHARMM22, AMBER99sb, and with the MP2/6-311(1d,1p) quantum basis set. The computations were done on the tripeptide *in vacuo* with the permutations of  $\phi$ ,  $\psi$  angles applied to all three Ala residues simultaneously.

### Analysis of protein helices by nominal crystallographic resolution

Analysis of all experimental structural models of helices in proteins in the Protein DataBank (Tables 1 and 2) exposed subtle inconsistencies in the torsional angle parameters of experimental high-resolution models compared to the classic  $\alpha$ - and  $3_{10}$ -helical models [37]. At high resolutions, where the experimental electron density accurately constrained helical conformations with the need for conformational assumptions and/or force field constraints, the backbone torsional angles populated a smooth distribution of helical conformations with a single maximum at  $\phi = -62$ ,  $\phi = -43$  (Fig 4). This is in contrast to the anticipated bimodal distribution that would result from a mixture of classic  $\alpha$ - and  $3_{10}$ -helices. Hovmoller et al. previously analyzed a non-redundant set of 1042 protein subunits from the PDB of Jan-2002; each structure was determined by X-ray diffraction to a resolution of  $2.0\text{ \AA}$  or better, refined to  $R \leq 0.20$ , and had less than 30% sequence homology [38].

The historical use of monopole electrostatics in force fields used to build atomic models from lower-resolution electron density appears (Fig 5) to have biased the PDB with early protein structures in which the presence of linear hydrogen bonding is over-represented. While

the magnitude of changes in  $\phi, \psi$  backbone torsion angles required to accurately fit high-resolution crystallographic data is small (i.e., only a few degrees), these errors can have a dramatic impact on other aspects of atomic structural models, from helical parameters propagating error at each helical turn, to significant errors in dipole moments, and most critically, upon the thermodynamics of interaction with helical surfaces. In order to address these issues, the geometrical relationships (Eqs 2 & 3) between backbone torsion angles and helical parameters that impact  $C\alpha - C\beta$  vectors were revisited.

## Impact of helical parameters

The almost unanimous support for the two classical helices by protein crystallographers over several decades was a serious concern. If backbone angles were really different and a strongly coupled relationship existed between helical parameters and backbone torsional angles, then changes in helical parameters should have been detected many times; if not in early crystal structures, then surely in the many tens of thousands of models that have been validated since! Starting with fundamental definitions for helical conformations, the effect of observed changes as a function of backbone torsional angles were plotted in terms of number-of-residues-per-turn ( $n$ ) and rise-per-residue ( $p$ ) (Fig 6). The implications of Figs 7, 8, 9 and 10 are profound, and may help explain why the classical  $\alpha$ -helix has enjoyed ongoing support in the crystallographic community.

The helical parameters ( $n, p$ ) for the experimental Némethy-, N- or  $3.6_{13/10}$ -helix are identical to those of the classical  $\alpha$ -helix, and unless the crystallographic resolution for a particular experiment is exceptional (i.e.,  $< 1.5 \text{ \AA}$ ), then the electron density associated with the carbonyl oxygen of the amide was insufficient to discriminate between helical structures. The use of inaccurate electrostatic treatments (i.e., monopole force fields), has also been a confounding issue, due to historical refinement methods that employing these potentials. In other words, the electron density from the heavy atoms of the peptide backbone (i.e., carbonyl carbon, amide nitrogen,  $C\alpha$ , and  $C\beta$ ) occupied identical positions in space for the classical  $\alpha$ -helix and the experimental Némethy-, N- or  $3.6_{13/10}$ -helix. Only the highest-resolution structures would be expected to discriminate between the two based only upon the position of the carbonyl oxygen, and such differences would be subtle. Initially, observations of the high-resolution helical population were assumed to impact the overall helical parameters, which seemed reasonable as the high-resolution population is intermediate between the  $\alpha$ - and  $3_{10}$ -helical populations and the  $3_{10}$ -helix is longer and tighter than the  $\alpha$ -helix. However, theoretical analysis of the dependence of helical parameters on torsional angles (Figs 7, 8, 9, and 10) resolved this apparent dilemma.

A compensatory relationship between  $\phi$  and  $\psi$  due to a crankshaft-like motion [39] of the planar amide bond between the two torsions allows orientation of the carbonyl oxygen semi-independently from the height, pitch, and width of the helix. To accommodate the rotation of the amide bond about an axis connecting the two  $\alpha$ -carbons (crank-shaft rotation), minor adjustments to bond angles are required, but energetic constraints on bond angles are comparatively low. This rationalizes the observations of three-centered helical hydrogen bonds in high-resolution protein structures that have been previously annotated/modeled as  $\alpha$ -helices [9].

## Comparison with protein helices from high-resolution crystal structures

Protein helices from experimental data with the most precise X-ray determination of atomic positions supports the analysis of Némethy-, N- or  $3.6_{13/10}$ -helices. The impact of torsion parameters on helix geometry can be independently accessed because the abundance of observations ensures that a very high-resolution structural model is built from experimental data with

**Table 4. Backbone Torsional Angles of High-resolution (1ejg, 0.54 Å) Structure of Crambin 7–18 Helix (Fig 11).**

| Residue # | Amino Acid | $\Phi$ | $\Psi$ | $\Phi$ Deviation | $\Psi$ Deviation |
|-----------|------------|--------|--------|------------------|------------------|
| 7         | Ile        | -52.6  | -53.6  | 7.9              | -11.7            |
| 8         | Val        | -53.2  | -49.3  | 7.3              | -7.4             |
| 9         | Ala        | -61.1  | -42.0  | -0.6             | -0.0             |
| 10        | Arg        | -63.0  | -44.6  | -2.6             | -2.7             |
| 11        | Ser        | -60.1  | -42.1  | 0.4              | -0.2             |
| 12        | Asn        | -60.7  | -43.8  | -0.3             | -1.9             |
| 13        | Phe        | -59.7  | -45.6  | 0.8              | -3.7             |
|           | Average    | -60.5  | -41.9  |                  |                  |
|           | N-Helix    | -62    | -43    |                  |                  |

doi:10.1371/journal.pone.0123146.t004

minimal contribution from geometric/energetic constraints. The small protein crambin with 46-residues has helices from experimental data obtained at very high resolution (PDB = 1ejg, 0.54 Å)[59]. A more highly refined crambin structure (PDB = 3nir, 0.48 Å) has recently been reported [60]. In Table 4 are listed the experimental  $\Phi, \Psi$  torsion values for helix (7–18) of crambin 1ejg (Fig 11). While the helical structure is obvious, the backbone torsion angles of the helical residues show considerable variation implying that they are sensitive to their environment, once again supporting the analysis of Hollingsworth et al. that secondary structures in proteins are not linear groups [32]; i.e., protein secondary structures do not behave as homopolymers with identical conformational parameters per residue. The shared hydrogen bridges in helices facilitate environmental adjustment by reducing the energetic barriers to optimize backbone torsional angles for side-chain interactions, and to respond to the local electrostatic environment [61]. In fact, MD simulations indicate that even short homopolymers exhibit dynamics heterogeneity in conformational space [62].

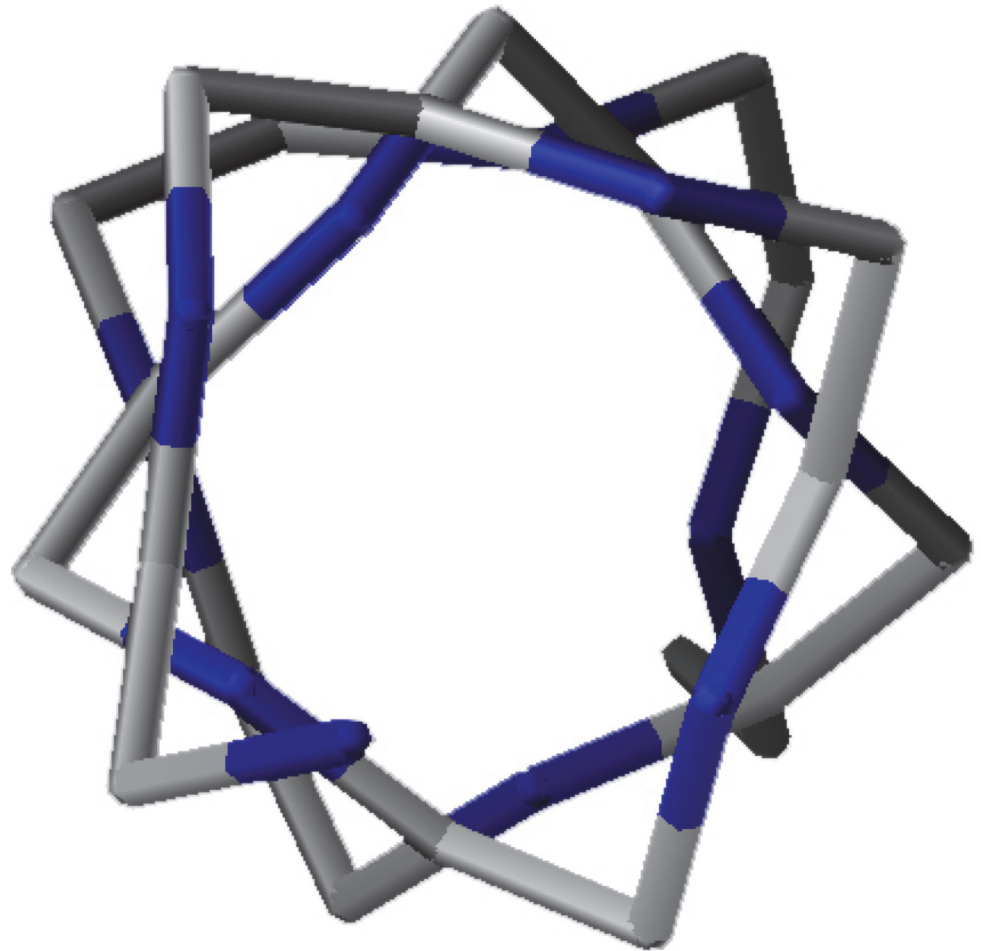
### Analysis of molecular dynamics simulations

As the structure was started in a helical conformation, this was confirmed by monitoring the changes in conformation as a function of length of simulation (Fig 12). A Ramachandran plot from the MD simulation demonstrated significant sampling of torsional space different from the starting helical conformation to provide evidence for sufficient sampling (Fig 13). Only a single peak was observed in the region of  $\phi = [-100 \rightarrow 0]$ ,  $\psi = [-70 \rightarrow 0]$  region. To further determine the center of the structures, structural filtering based on  $\alpha$ - and  $3_{10}$ -helical hydrogen bond lengths (i.e.,  $< 4$  Å) and  $\phi, \psi$  angles was performed (Fig 14). Both  $\phi, \psi$  angles and  $i \rightarrow i+4$  distances followed normal distribution with slight skewing to the left of  $i \rightarrow i+4$  hydrogen bonds. Considering the strong penalty for steric overlap, an approximately normal distribution resulted. The “average” structure had helical backbone torsions  $\phi = -72.78$ ,  $\psi = -33.43$  and  $i \rightarrow i+3 = 2.97$  Å;  $i \rightarrow i+4 = 2.28$  Å hydrogen-bond distances (Fig 15) with a conformation that resembled the  $\alpha$ - and experimental Némethy-, N- or  $3.6_{13/10}$ -helix (Fig 16).

More detailed analyses of hydrogen-bonding dynamics from an oligoalanine MD simulation with AMOEBA can be found in a paper by Liu et al. [62] comparing the relative stabilities of helices comprised of capped 12-residues peptides of oligoglycine, oligoalanine and oligo- $\beta$ -alanine.

### Quantum calculations on Ac-Ala-Ala-Ala-Nme

In order to observe any significant electronic differences between  $\alpha$ -, experimental Némethy-, N- or  $3.6_{13/10}$ - and this dynamic-helix, MP2/6-311(1d,1p) quantum calculations were



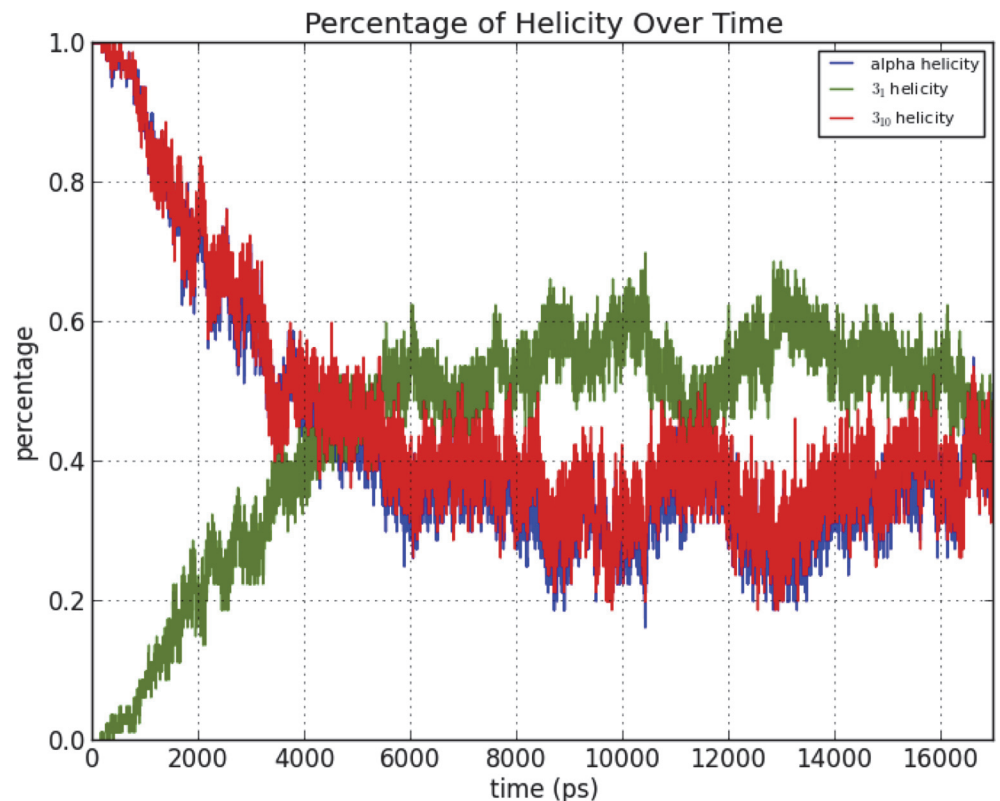
## Crambin Helix Val 8 to Leu 18

**Fig 11. Helix 8–18 from the crambin crystal structure [PDB:1ejg], viewed from above the N-terminus.** The average  $\phi$  value along the helix was between the experimental Némethy-, N- or  $3.6_{13/10}$ -helix and the  $\alpha$ -helix. However, the average  $\psi$  value was between the experimental  $3.6_{13/10}$ -helix and a  $3_{10}$ -helix.

doi:10.1371/journal.pone.0123146.g011

performed on a simple capped tripeptide model, Ac-Ala-Ala-Ala-NMe, first *in vacuo*, and then with the PCM hydration model [63]. The relative energies and minima locations of the three helical conformers are given in Tables 5 & 6. The electron densities of the experimental Némethy-, N- or  $3.6_{13/10}$ - and dynamic-helices were very similar. But in the classic  $\alpha$ -helix, HN39 was slightly more polarized with its electron density oriented towards O3.

A large decrease in energy for the  $\alpha$ -helical region was observed by the PCM calculation, which lowered its minimum to 0.45 kcal/mol lower than the  $3_{10}$ -helix (approximately a 2:1 ratio at room temperature) with an energy barrier of 0.7 kcal/mol ( $\sim kT$ ) between the  $3_{10}$  and  $\alpha$ -helical regions. This can be explained by the interaction energy of PCM water with the hydrophilic groups on the capped tripeptide that were not hydrogen bonded, and thus a compensation of the strong *in vacuo* end effects. But besides this effect, the center of the  $3_{10}$ -helical well also shifted from (-70, -15) to (-65, -20), a more depolarized, shorter  $i \rightarrow i+3$  distance and similar  $i \rightarrow i+4$  distance structure. The  $\alpha$ -helical well was now centered about (-65, -37). Both



**Fig 12. Transition of oligoalanine from its starting  $\alpha$ -helical conformation to intermediate states that includes near  $\alpha$ -,  $3_1$ - and polyglycine conformations.**

doi:10.1371/journal.pone.0123146.g012

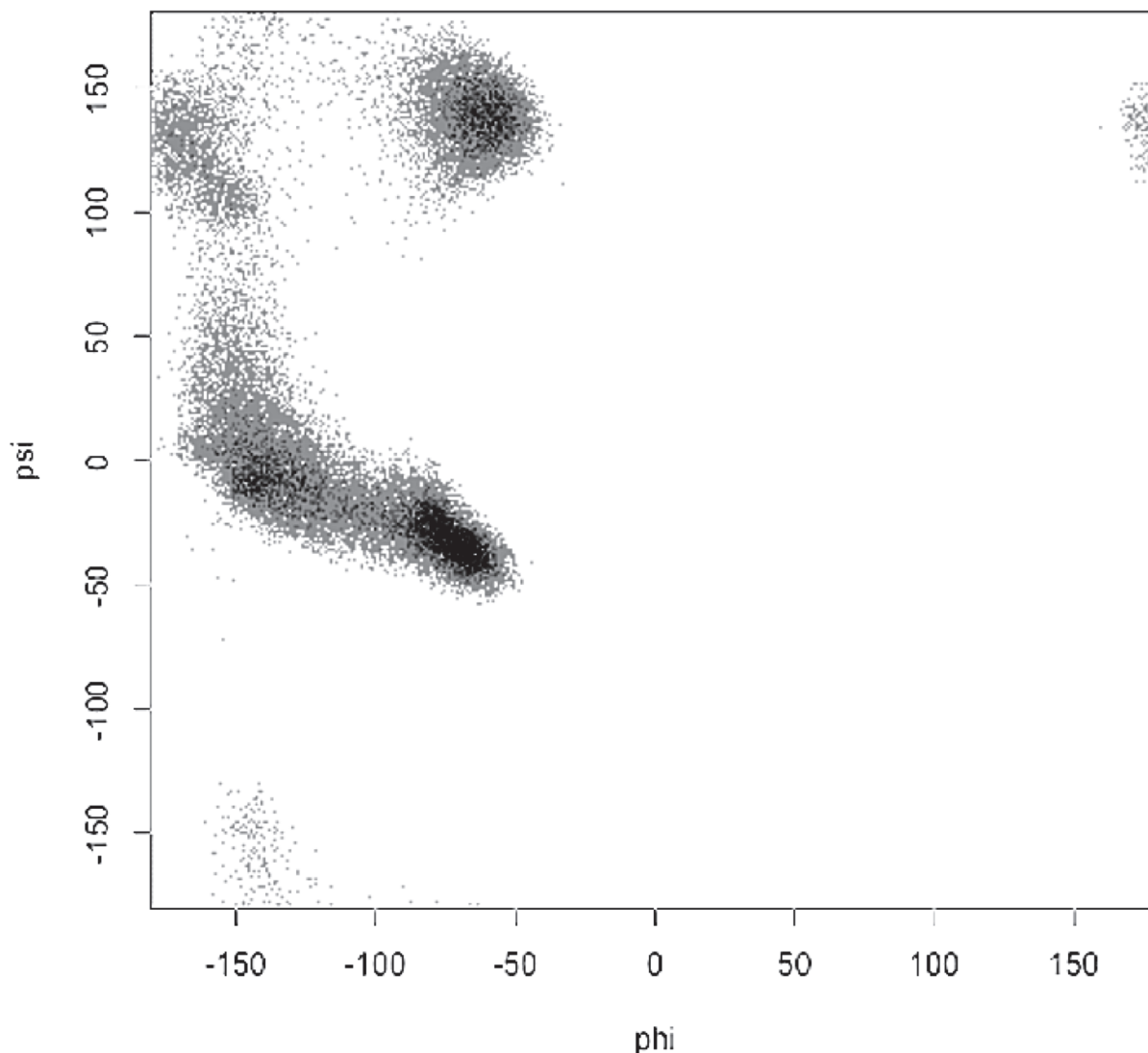
centers shifted to tighter and more polarized structures than *in vacuo*. It is well-known that stability of  $\alpha$ - versus  $3_{10}$ -helices in short peptides depends strongly on length and dielectric of the solvent [64]. This arises primarily from the extra hydrogen bond in a classical  $3_{10}$ -helix versus an  $\alpha$ -helix of the same length. Further QM calculations on peptides of different lengths are required to resolve the end effects seen with the tripeptide.

### Estimated dipole moments for different helical conformations

To compensate for the strong end effects of the short tripeptide model, capped twelve-residue oligoalanines were generated with backbone torsion angles set to those of  $\alpha$ -helix,  $3_{10}$ -helix, experimental Némethy-, N- or  $3.6_{13/10}$ -helix, the MD helix and the QM helix. PDB files of these structures were submitted to a web-based server for Adaptive Poisson-Boltzman Solver described by Felder et al. [65] to calculate protein dipoles and the results are presented in Table 7.

The experimental  $3.6_{13/10}$ -helix has a calculated dipole moment most similar to the  $3_{10}$ -helix reflecting the reorientation of the amide bond from that of the classic  $\alpha$ -helix. Sengupta et al. reported a calculated dipole moment of 42.7 Debye for a oligoalanine  $\alpha$ -helix of twelve residues using CHARMM monopole charges [66]. For comparison, TINKER was used to calculate the dipole moments using AMOEBA multipole charges. The differences in magnitude (Table 7) are significant, and the relative values change as well with a maximum difference of 7 Debye (74–67) for the dipoles from the server, and 10.2 Debye (62.1–51.9) for AMOEBA dipoles. Obviously, the redistribution of charge using multipoles versus monopoles significantly impacts the estimation of dipole moments. In order to further verify significant differences



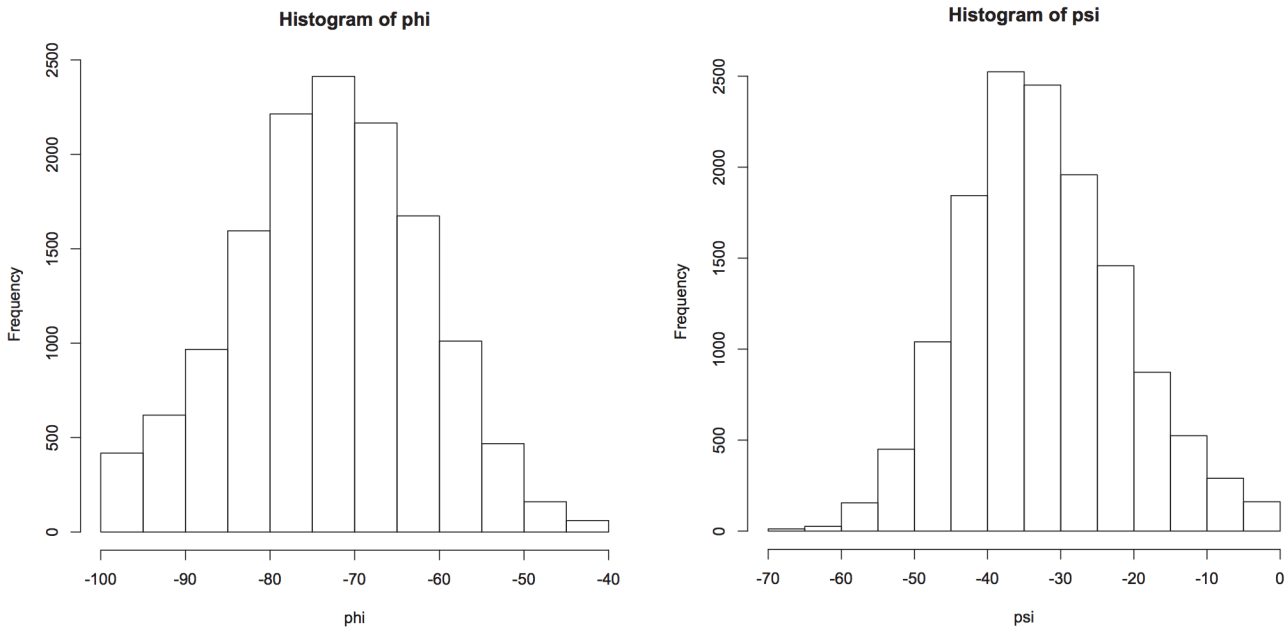


**Fig 13. Ramachandran plot of backbone torsional angles from the AMOEBA BIO09 MD simulation of 12-residue oligoalanine.**

doi:10.1371/journal.pone.0123146.g013

between the approaches, dipole moments were calculated for crambin (PDB: 3nir, 0.48 Å) and its isolated helix (Ac-6-19-NHMe). The switch in relative magnitudes between dipole moments calculated with a monopole approximation versus a multipole method for crambin and its isolated helix ([Table 7](#)) dramatically emphasize the importance of multipoles in electrostatic interactions.

Ripoll et al. analyzing the alignment of peptide-bond dipole moments with the local electrostatic field generated by the rest of the molecule for a large set of native proteins; dipole moments associated with  $\alpha$ -helical conformations show the best alignment with the electrostatic field [68]. The large difference in dipole moment seen for the different helical conformations suggests that the electrostatic environment adjacent to a protein helix could modulate its backbone torsional angles with minimal perturbation to side-chain orientations and their steric interactions.

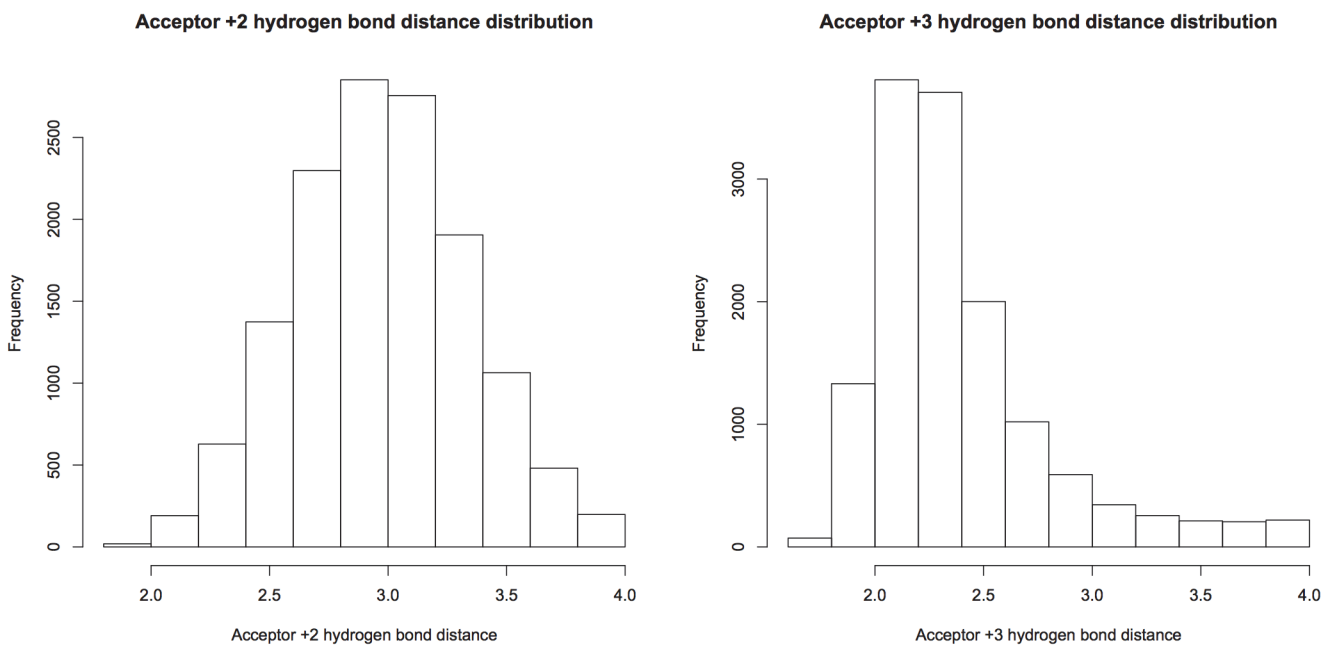


**Fig 14. The  $\Phi$ ,  $\Psi$  angles distribution of the data filtered with H-bond distance ( $<4 \text{ \AA}$ ) and  $\Phi$ ,  $\Psi$  angles range ( $-100 \rightarrow 0$ ,  $-70 \rightarrow 0$ ).** The median of  $\Phi$ ,  $\Psi$  angles was  $\Phi$ -72.8,  $\Psi$ -33.4.

doi:10.1371/journal.pone.0123146.g014

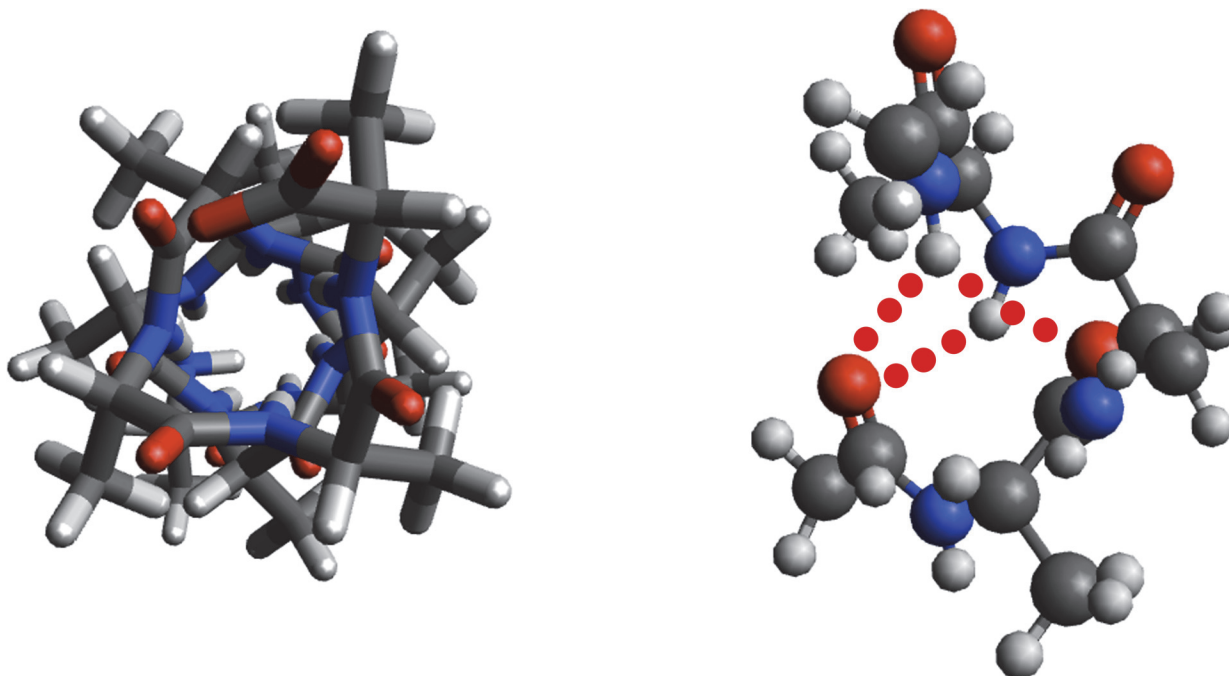
### *In vacuo* comparisons of helical minima and activation energies for transitions

As a baseline for comparison, the potential surfaces around helical conformations of Ac-Ala-Ala-NMe were generated *in vacuo* using the three monopole force fields and AMOEBA.



**Fig 15. Distance distribution of observed H-bonds filtered by H-bond distance ( $<4 \text{ \AA}$ ) and  $\Phi$ ,  $\Psi$  angles range ( $-100 \rightarrow 0$ ,  $-70 \rightarrow 0$ ).** The medians of  $i+3$  and  $i+4$  distances were 2.97 Å and 2.28 Å, respectively.

doi:10.1371/journal.pone.0123146.g015



**Fig 16. Orthogonal views of ball-and-stick model of “dynamic” helix of 12-residue capped oligoalanine with two three-centered hydrogen bonds indicated (red dots).**

doi:10.1371/journal.pone.0123146.g016

These were compared with MP2/6-311(1d,1p) quantum calculations that sampled the same grid points ( $\phi = -100 \rightarrow 0$ ,  $\psi = -70 \rightarrow 0$ ). Energies greater than the minimum by more than 20 kcal/mol were normalized to make the picture clearer. The six plots (Fig 17) show potential surfaces around the helical conformers, which clearly demonstrate significant differences in conformer stability and the limited activation energy barriers between them.

To further characterize these potential surfaces, the locations of the activation energy required for transitions between minima were tabulated. On the contour plots, both QM and the monopole force fields show potential funnels with two minima-like positions. From the contour plots for all the potential surfaces, no energy barrier existed between the two minima, and a “U” shape path would lead the structure from the higher energy position to the lower energy one without any activation-energy barrier (Fig 18). By QM calculations, the minimum energy-like conformer by QM was located at  $\phi = -70$ ,  $\psi = -17$ , while the other conformer was found at  $\phi = -73$ ,  $\psi = -33$ . The first was close to the classic  $3_{10}$ -helix and the second closer to the “dynamic” helix. Thus, they approximate the  $3_{10}$  and  $\alpha$ -helix regions, respectively. This suggests that the  $3_{10}$ -helix was energetically the most stable structure for Ac-Ala-Ala-Ala-NHMe *in vacuo*. This is plausible since the more compact hydrogen-bonding pattern of the  $3_{10}$ -helix allows one more hydrogen bond than the  $\alpha$ -helix for the same length of peptide [64].

**Table 5. Relative energies of the three helical structures of Ac-Ala-Ala-Ala-NHMe by MP2/6–311 (1d,1p) quantum calculations *in vacuo*.**

|                                      | Energy/Eh | Relative Energy/ kcal |
|--------------------------------------|-----------|-----------------------|
| classic $\alpha$ -helix              | -988.0102 | 1.958                 |
| Némethy-, N- or $3.6_{13/10}$ -helix | -988.0134 | 0.00                  |
| “dynamic”                            | -988.0126 | 0.501                 |

doi:10.1371/journal.pone.0123146.t005

**Table 6. Location of the two energy minima-like positions on the potential energy surfaces of Ac-Ala-Ala-Ala-NHMe based on methodology.**

|                  | Position of the minimum  | Position of the second funnel |
|------------------|--------------------------|-------------------------------|
| MP2/6-311(1d,1p) | $\Phi = -70, \Psi = -17$ | $\Phi = -73, \Psi = -33$      |
| Amber99sb        | $\Phi = -69, \Psi = -16$ | $\Phi = -71, \Psi = -34$      |
| AMOEBA BIO09     | $\Phi = -69, \Psi = -37$ | N/A                           |
| CHARMM22         | $\Phi = -75, \Psi = -30$ | $\Phi = -78, \Psi = -20$      |
| OPLS-AA          | $\Phi = -74, \Psi = -33$ | $\Phi = -71, \Psi = -16$      |
| OPLS-AAL         | $\Phi = -75, \Psi = -30$ | $\Phi = -79, \Psi = -11$      |

doi:10.1371/journal.pone.0123146.t006

The QM potential well (Fig 18) supports the absence of two canonical helices. Whereas there appear to be two minima on the QM plot, the activation energy barrier between them, if any, was so low that they are functionally merged. All of the calculated potential surfaces (Fig 17) have transitions between the two detected minima similar to that of the QM surface. Thus, none of the potentials required a peptide to overcome an activation barrier between helical minima supporting the conclusion that there is no basis for the classical  $\alpha$ - and  $3_{10}$ -helix dichotomy, and supporting the hypothesis of the “molten helix” suggested in 1995 by Smythe et al. [69,70]. In addition, the AMOEBA surface (Fig 17(B)) most resembles the QM surface (Fig 17(F)) in depth of the potential well, but differs slightly in absence of distinct minima corresponding to the two helical conformations. Despite direct parameterization using QM, the AMOEBA potential surface does not exactly duplicate that from QM suggesting that further improvements in parameterization can still be made.

### Evidence from high-resolution crystal structures of peptides supporting $\alpha$ - and $3_{10}$ -helices

Abundant high-resolution crystal structures of small peptides clearly show the two classical helices [71–73] as well as transitions between them depending on sequence [74,75] and crystallization conditions [76,77]. In most of these cases, the crystalline peptides contained one, or more, *gem*-dimethyl amino acids, such as  $\alpha$ -aminoisobutyric acid (AIB,  $\alpha$ -methylalanine) or another amino acid in which the  $\alpha$ -proton was replaced by a methyl group. Aib is highly heliogenic and surveys of synthetic Aib-containing peptides show a preponderance of  $3_{10}$ -, -, and mixed  $3_{10}/\alpha$ -helical structures. Aib residues have been observed in non-helical conformations, primarily into the polyproline II (P(II)) and fully extended regions of conformational space [78]. The achiral Aib residue can adopt both left ( $\alpha(L)$ )- and right ( $\alpha(R)$ )-handed helical conformations. These experimental observations confirm the theoretical predictions of Marshall and Bosshard on the impact of methyl for proton substitution on peptide backbones made in 1972 [73].

The resolution obtained by direct methods on small molecules exceeds that normally obtained for proteins. A strong argument can be made that the data from Aib-containing helical

**Table 7. Calculated dipole moments [65] in Debye using a modified set of parameters for solvation energy (Parse) partial atomic charges [67] for various capped 12-residue oligoalanine helices as well as for crambin and its isolated helix compared with those using AMOEBA multipoles.**

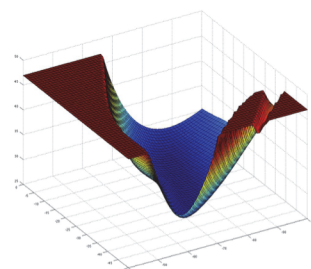
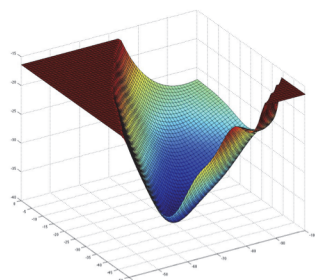
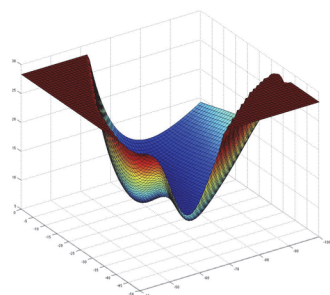
|                                      |             |                              |             |
|--------------------------------------|-------------|------------------------------|-------------|
| $\alpha$ -helix                      | 74 vs. 60.2 | $3_{10}$ -helix              | 70 vs. 62.1 |
| Némethy-, N- or $3.6_{13/10}$ -helix | 70 vs. 62.1 | MD-helix                     | 72 vs. 55.4 |
| A-Helix                              | 72 vs. 55.6 | QM-helix                     | 67 vs. 51.9 |
| Crambin                              | 54 vs. 64.6 | Crambin helix (Ac-6-19-NHMe) | 88 vs. 70.9 |

doi:10.1371/journal.pone.0123146.t007

A. Amber 99s

B. AMOEBA

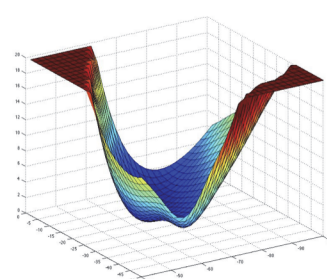
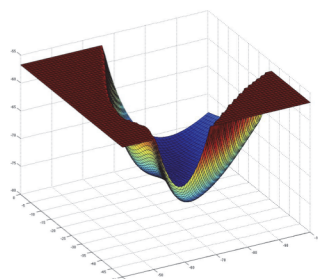
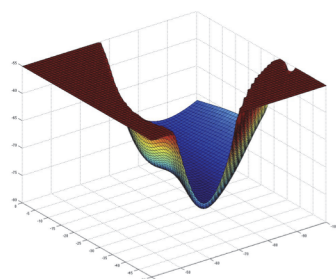
C. CHARMM22



D. OPLS-AA

E. OPLS-AAL

F. QM



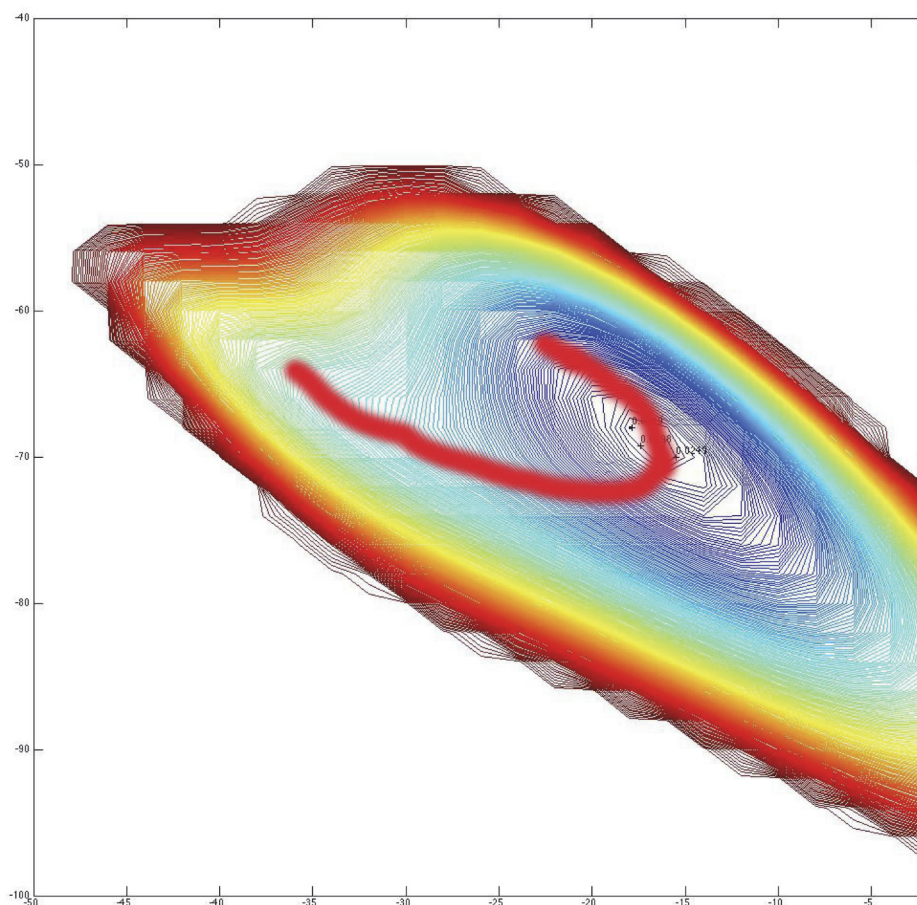
**Fig 17. Potential surfaces for torsional force-fields surfaces surrounding helical conformations for A. AMBER99sb, B. AMOEBA, C. CHARMM22, D. OPLS-AA, and E. OPLS-AAL versus F. QM—MP2/6-311(1d,1p) basis set.**

doi:10.1371/journal.pone.0123146.g017

peptides overrules the rejection of the two classical helices based on data from the PDB. To quote from an early study of Aib oligomers by Paterson et al. [79], “Tetrahedral symmetrical geometry for these substituents favors the  $\alpha$ -helical conformation for the  $\alpha$ -aminoisobutyric acid residue (and di- and tripeptides thereof) whereas asymmetric geometry, derived from well-refined X-ray structures, gives the  $3_{10}$  conformation as the preferred structure. Analysis of pairwise atomic interactions indicates that favorable backbone-backbone interactions lower the overall energy of the molecule in the  $3_{10}$  conformation when the substitution on the  $C^\alpha$  atom is asymmetric.”

There is, however, a plausible explanation of the observations seen with Aib-containing helical peptides; the presence of the *gem*-dimethyl groups on the peptide backbone favor the  $3_{10}$ -helix due to steric interactions [80]. In fact, Hodgkin et al. proposed a  $3.6_{10}$ -helix in 1990 [80] based on the structure and hydrogen-bonding pattern of Tos-(Aib)<sub>5</sub>-OMe [81] that existed in a helix with  $n$  to  $n-3$  H-bonding and an angle of  $\sim 100$  degrees between  $\alpha$ -carbons, similar to the  $3.6_{13/10}$  proposed herein based on high-resolution crystal structures. This helix was very similar to one proposed earlier by Némethy et al. in 1967 [9] based on analytical analysis of helical parameters, except the  $\alpha_{II}$ -helix had no hydrogen bonding. An intermediary helix between the  $\alpha_I$ - and  $\alpha_{II}$ -helix shown in the paper (Fig 3), however, had “hydrogen bonding arrangements of these residues is intermediate between those of the  $\alpha_I$ -helix and the  $(3_{10})_I$ -helix, while the helical parameters  $n$  and  $h$  are still those for the  $\alpha_I$ -helix.” [9] It is ironic that some 47-years later, the “Némethy”-helix ( $3.6_{13/10}$ -helix) has been validated as dominant in protein structures, and





**Fig 18. The potential surface near helical torsional angles for Ac-Ala-Ala-Ala-NMe as calculated by QM.** The red line traces the transition between the two energy minima-like conformers without an activation energy barrier.

doi:10.1371/journal.pone.0123146.g018

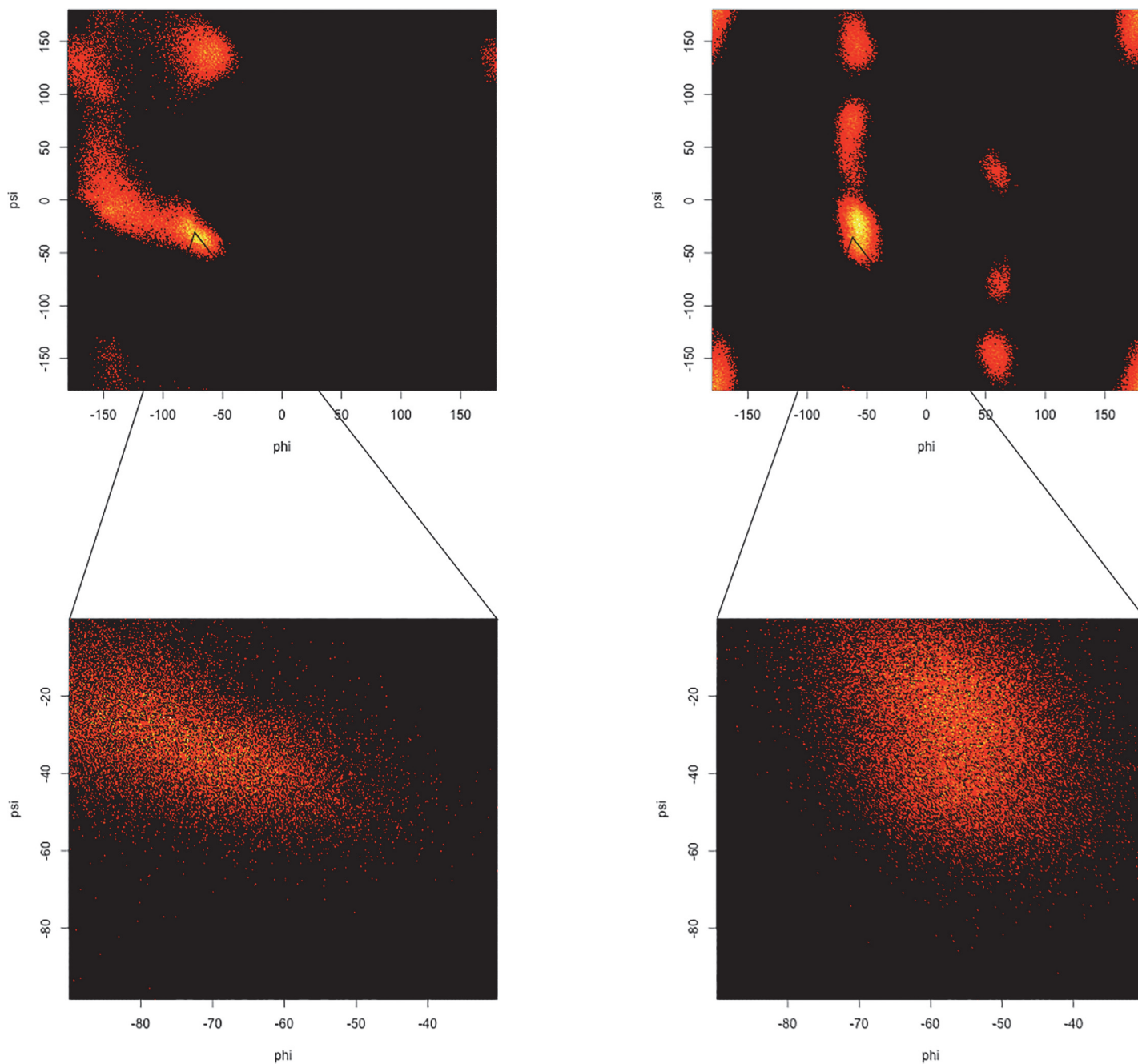
by a group including one of the students (GRM) whom Prof. George Némethy tutored at Rockefeller University.

While the analysis of Hodgkin et al. was very thorough, it depended on an early version of AMBER. In order to test this hypothesis with a more precise force field, MD simulations in water of 12-residue capped oligoalanine and oligo- $\alpha$ -aminoisobutyric acid were performed with AMOEBA. The Ramachandran plots from the 10-nsec replica-exchange MD simulations are compared (Fig 19). The hydrogen-bonding matrices for oligo-Ala and oligo-Aib from the MD simulations (Fig 20) show a significant increase in  $3_{10}$ -like helices in the oligo-Aib peptide. Thus, introduction of Aib residues enhances the preference for  $3_{10}$ -like helices and limits the relevance of data from such peptides to protein helices.

### Role of three-atom-centered hydrogen bonds

Hydrogen bonds are dominated by electrostatic forces, and are consequently weak, directional, promiscuous, with compensatory features not accurately represented by a single linear potential. Jeffrey and Saenger forecast in 1991, “In future high-resolution crystal structure studies of proteins, hydrogen bonding has to be analyzed in terms of two-centered and three-centered interactions. It should be stressed that three-centered bonds may be especially important in



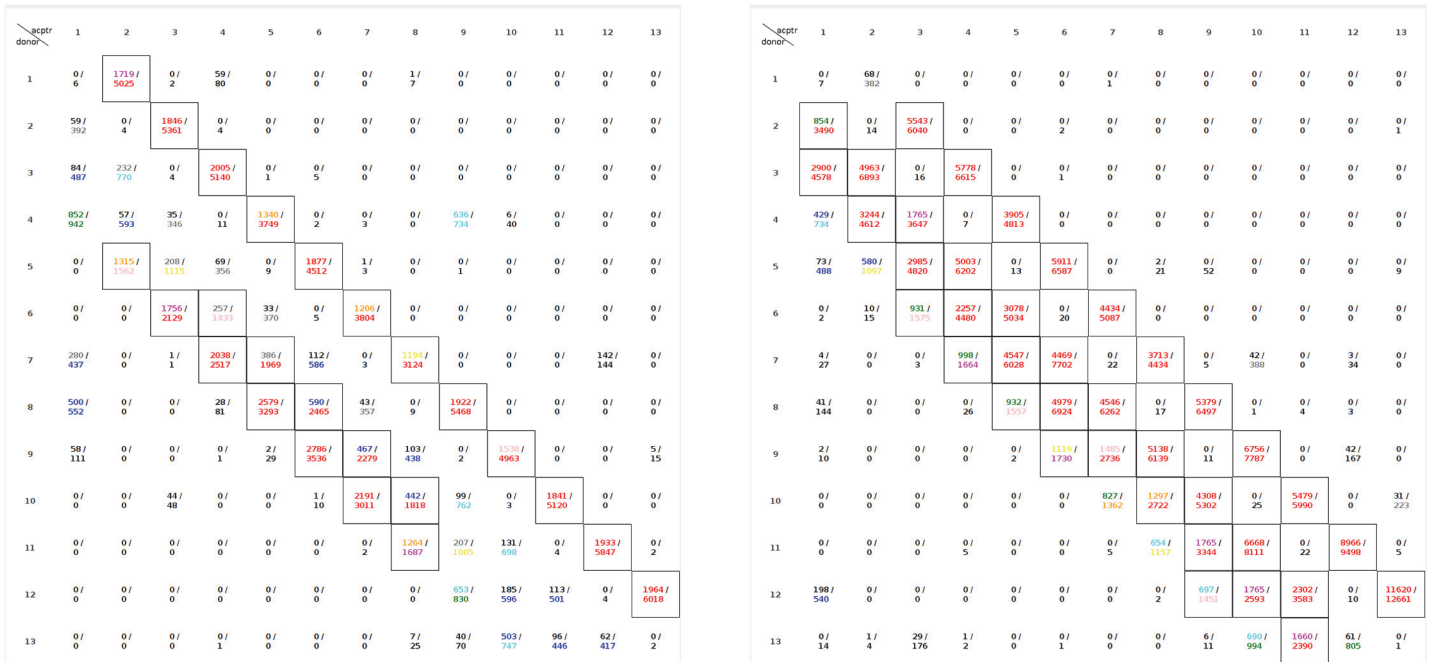


**Fig 19. The whole (top) and expanded (lower) Ramachandran plots of 12-residue oligo-Ala (left sides) and oligo-Aib (right sides) from AMOEBA replica-exchange MD simulation. The  $3_{10}$ -region in Aib is heavily occupied while in alanine it's almost absent.**

doi:10.1371/journal.pone.0123146.g019

protein dynamics because they can be regarded as transitional intermediates from one two-centered bond to another.” [6]

Simultaneous, shared hydrogen bonds are conceptually equivalent to the networked hydrogen bonds in proteins as recognized by Stickle et al. [41], “. . . almost all central residues of helices are within networks because helix geometry is commensurate with both  $i \rightarrow i+4$  and  $i \rightarrow i$



**Fig 20. The hydrogen-bonding matrix of 12-residue capped oligo-Ala (left) and oligo-Aib (right) from AMOEBA replica-exchange MD simulation.** The acceptor+1 ( $i \rightarrow i+2$ ) and acceptor+2 ( $i \rightarrow i+3$ ), indicating  $3_{10}$  helix) in oligo-Aib is more heavily occupied than in oligo-Ala.

doi:10.1371/journal.pone.0123146.g020

+3 hydrogen bonds simultaneously. . .” and “. . . it is likely that the existence of multiple determined hydrogen bonds enhances the elasticity of an  $\alpha$ -helix”

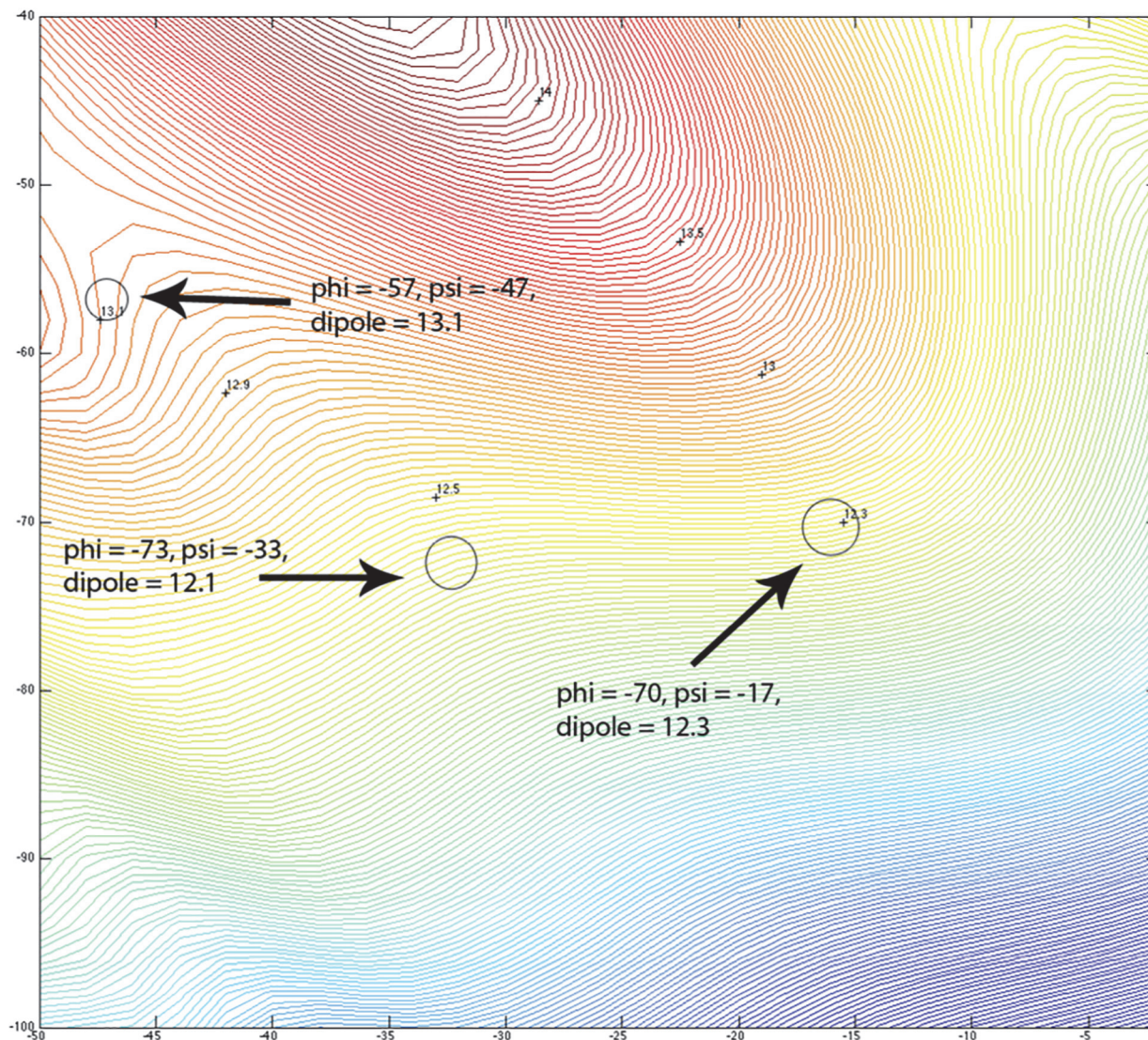
Mixed DFT/AM1 quantum calculation on oligoalanine helices by Wiczorek and Dannenberg [82] found “hybrid”  $3_{10}/\alpha$  structures in which “C = O forms both  $3_{10}$ - and  $\alpha$ -helical H-bonds.” Bour et al. found average torsional angles of  $\phi = -63.3$  and  $\psi = -42.1$  for capped alanine nonapeptides using HF/DFT with COSMO solvation [71]. Thus, a consensus from both experimental and theoretical methods strongly supports three-atom-centered (bifurcated) hydrogen bonds as found in the Némethy, N-,  $3.6_{13/10}$ -helix first proposed in 1967 [9].

Nevertheless, the classical  $\alpha$ -helical and  $3_{10}$ -helical models of linear hydrogen bonds still dominate structural literature on proteins, and strongly influence conceptual views of protein dynamics. Estimates of the strength of bifurcated hydrogen bonds by quantum mechanics in model systems indicate negative cooperativity regarding the calculated interactive energy per hydrogen bond [83,84]. In effect, donors and acceptors become less avid hydrogen-bonding partners experimentally once they are already participate in a hydrogen bond [85], but the total stabilization of the two hydrogen bonds in the three-centered hydrogen bond is greater than a single two-centered hydrogen bond. The actual energetics of bifurcated H-bonds in proteins are still under investigation [86].

### Monopole vs. Multipole Electrostatics

More than three decades of literature clearly cite the inability of monopole electrostatics to reproduce the molecular electrostatic potential calculated by *ab initio* quantum mechanics. For example, Williams showed that inclusion of dipole and quadrupole terms were required to accurately mimic *ab initio* electrostatic potentials of small molecules, such as water [20,21]. Stone has argued since the 1980s that distributed multipole analysis (DMA) derives a much more accurate description of the electrostatic potential surrounding a molecule in terms of charges,





**Fig 21. Contour plots of oligo-Ala dipole moment.** The dipole moment magnitudes of the classic  $\alpha$ -helix,  $3_{10}$ -helix and Nemethy-, N-,  $3.6_{13/10}$ -helix are marked with circles.

doi:10.1371/journal.pone.0123146.g021

dipoles, and quadrupoles compared with a calculated wave function [19,22,87]. Ren and Ponder have shown that inclusion of polarizability and multipole electrostatics, derived with DMA, was necessary to reproduce the colligative properties of bulk water and the local geometry of water clusters using a single consistent set of parameters [18]. Schneiders et al. have presented a convincing case for the use of the polarizable AMOEBA force field in crystallographic refinement [28]. In preliminary studies comparing molecular dynamics simulations with explicit solvent of high-resolution crystal structures using AMOEBA versus OPLS-AA and CHARMM, Kuster [37] found that AMOEBA preserved the experimental helical length of crambin 7–18 (PDB = 1CRN, resolution = 0.945 Å) while OPLS-AA and CHARMM both shortened the helix approximately 0.5 Å to a length corresponding to an ideal  $\alpha$ -helix. These results support the view that monopole electrostatics used in some low-resolution structural refinements introduced  $\alpha$ -helical bias.

## Dipoles of helices and external fields

In order to determine the dipole change during the transition from  $3_{10}$ -helix to  $\alpha$ -helix, the dipole moments from QM *in vacuo* in the  $\alpha$ -helical region was plotted (Fig 21). The reorientation of the amide bond seen in the experimental Némethy-, N- or  $3.6_{13/10}$ -helix might be expected to impact the dipole moment of the capped trialanine helix. Hol [88] and others [89,90] have considered the roles of helix dipoles in protein function. Ripoll et al. have shown that the dipole moments associated with  $\alpha$ -helical conformations showed alignment with the local electrostatic field in a large set of native proteins [68]. The QM-calculated dipole for the different Ac-Ala-Ala-Ala-NMe helices in this study varied from 12.1 Debye for the Némethy-, N-,  $3.6_{13/10}$ -helix to 13.1 Debye for the  $\alpha$ -helix (Fig 21), hardly a significant change.

Calculation of the dipole moments for Ac-(Ala)<sub>12</sub>-NHMe, however, using a web-based server for Adaptive Poisson-Boltzman Solver [65] or with the AMOEBA charges gave very significant changes in dipole moments depending on helical conformation (Table 6). This implies another parameter for optimization in protein folding, orientation of the peptide-bond dipoles in response to the local electrostatic field that is consistent with the lack of linear groups observed in experimental protein structures [32]. Choudhary and Raines have found that amino acids, which adopt the main chain dihedral angles of an  $\alpha$ -helix, display dramatic pyramidalization of carbon (C $\alpha$ ) toward the oxygen (O $\alpha$ -1) leading them to conclude that O $\alpha$ -1 and C $\alpha$  are linked by a partial covalent bond in  $\alpha$ -helices [61]. This potentially significant quantum interaction has not been considered in this manuscript. Further implications of the reorientation of the amide bond remain to be explored. The current second-generation force field AMOEBA, however, appears adequately parameterized to both reproduce this phenomena, and to allow exploration of its impact on protein structures.

## Conclusions

The structural and energetics of the helical region of proteins has been examined by X-ray and neutron crystallography, molecular dynamics simulations, monopole and second-generation force fields that include multipole electrostatics and polarizability as well as quantum mechanics (focused primarily on the ( $\varphi = -100 \rightarrow 40$ ,  $\psi = -50 \rightarrow 0$ ) region of backbone torsional space). From a detailed examination of X-ray crystallography data in the PDB, systematic differences in torsion-angle distributions between high- and low-resolution experimental data were observed. A single Gaussian-like peak was observed in the high-resolution data, challenging the classical  $\alpha$ - and  $3_{10}$ -helical hypotheses that were based on linear hydrogen bonding. Since low-resolution X-ray crystallography and NMR data have utilized force-field constraints in refinement (until very recently, with inaccurate monopole electrostatics), serious concerns are raised regarding the accuracy of some data, particularly models derived for low-resolution diffraction data, in the Protein Data Bank.

The torsional distribution of oligoalanine in AMOEBA09 MD simulations in water showed a single peak at  $\varphi = -72.78$ ,  $\psi = -33.43$ , once again challenging the two canonical helices hypothesis resulting from monopole electrostatics and linear hydrogen bonding. The role of water in protein structure and dynamics cannot be overemphasized [91,92], and MD simulations, assuming an accurate force field, are a viable way to explore solvation. Overall, protein helices are dynamic and flexible in responding to their environments; only in unusual circumstances could they be considered “linear groups”. Rather, carbonyl and amide bond groups optimize the electrostatic interaction of their individual dipoles with the local, surrounding fields [68] which is generally possible to accomplish without requiring significant changes in the position or orientations of the peptide C $\alpha$ -C $\beta$  side chain vectors that impact packing. This is

consistent with the variation in backbone torsional parameters seen with the crambin helix ([Table 4](#)).

In order to reconcile the experimental and theoretical data with the geometrical requirements of molecular orbital theory and a modern treatment of electrostatics, it was necessary to introduce slight adjustments of the backbone torsional angles associated with protein helices while still maintaining the overall  $\alpha$ -helical parameters (rise-per-residue and pitch). Surprisingly, this crankshaft-like rotation of the amide bond does not move the  $C\alpha$  and  $C\beta$  atoms from their positions in the Pauling  $\alpha$ -helix, a fact that may explain ongoing structural support for linear hydrogen bonding. It is important to recognize that this crankshaft rotation does, however, accommodate three-centered amide hydrogen bonding (typically with a shorter and stronger  $i \rightarrow i+4$  compared to the longer and weaker  $i \rightarrow i+3$  hydrogen bond). Thus, minor refinement of helical backbone torsional angles to those of the consensus experimental high-resolution crystal structure ( $\varphi = -62$ ,  $\psi = -43$ ) is recommended as being most consistent with all available experimental and theoretical considerations. We designate this the Némethy-, or N-helix ( $3.6_{13/10}$ ) having the same helical parameters (rise-per-residue and pitch) as the classical  $\alpha$ -helix, and modified backbone torsional angles to accommodate three-centered hydrogen bonds.

The resulting perspective reconciles requirements of carbonyl molecular orbitals, the enhanced strength and entropic benefits of three-centered hydrogen bonds, second-generation force fields with multipole electrostatics and polarizability with high-resolution crystal structures of protein helices. This self-consistent view of protein helices abandons the linear hydrogen bond without modifying the overall (crystallographically observable) heavy-atom positions (with the exception of the oxygen of the amide carbonyl) and helical parameters of the classic  $\alpha$ -helix of Pauling, Corey and Branson [1].

## Acknowledgments

The authors appreciate early critical discussions with Prof. Anthony Stone of Cambridge University. Assistance in modeling helix dipoles by Ms. Claire Simons (Washington University) and with comparative minimizations of crambin by Dr. Sergio Urahata (Universidade de São Paulo, Brazil) is gratefully recognized. Mr. Andrew Meiburg and Prof. David H. Freemont of Washington University School of Medicine provided [Fig 4](#) that simulated that of Wlodawer et al. [33] when FASEB Journal would not allow its reproduction. Mr. Xin Gao (Washington University) helped derive the AMOEBA parameters for Aib. Our thanks to Profs. David Eisenberg of UCLA, Alfred Holtzer of Washington University, and Harold Scheraga of Cornell University for constructive comments.

This work is dedicated to the memory of György (George) Némethy, who was born in Budapest in 1934 and studied in Hungary and Germany before immigrating to the United States. He received his Ph.D. from Cornell University in 1962, having been mentored by Prof. Harold Scheraga, a future collaborator for his entire career. George Némethy was an assistant professor at Rockefeller University when his paper with Scheraga, “Theoretical determination of sterically allowed conformations of a polypeptide chain by a computer method,” was published in 1965 in *Biopolymers*. As GRM was a graduate student with Bruce Merrifield at the time working on the synthesis of peptides, this paper convinced him that computational chemistry was essential to his future research. Prof. Némethy agreed to tutor him in the mathematics of conformational enumeration, but only after much harassment. Thus, GRM indirectly joined the computational biophysics community as one more scientific disciple of Prof. Harold Scheraga.

George Némethy made many outstanding scientific contributions in addition to his helix predictions, producing close to 300 scientific publications. For his graduate work on



hydrophobic interactions in hydrogen-bonding solvents, such as water, he received the Pius XI Gold Medal of the Pontifical Academy of Sciences in 1972. Another insightful contribution was the KNF model of allostery published with Koshland and Filmer in 1966. He died of cancer in 1994 while Professor of Biomathematical Sciences at the Mount Sinai School of Medicine.

## Author Contributions

Conceived and designed the experiments: DJK CL ZF JWP GRM. Performed the experiments: DJK CL ZF. Analyzed the data: DJK CL ZF JWP GRM. Contributed reagents/materials/analysis tools: DJK CL ZF JWP. Wrote the paper: DJK CL JWP GRM.

## References

1. Pauling L, Corey RB, Branson HR (1951) The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. *Proc Natl Acad Sci U S A* 37: 205–211. PMID: [14816373](#)
2. Eisenberg D (2003) The discovery of the alpha-helix and beta-sheet, the principal structural features of proteins. *Proc Natl Acad Sci U S A* 100: 11207–11210. PMID: [12966187](#)
3. Pauling L, Corey RB (1950) Two Hydrogen-Bonded Spiral Configurations of the Polypeptide Chain. *J Am Chem Soc*: 5349.
4. (1970) IUPAC-IUB Commission on Biochemical Nomenclature. Abbreviations and symbols for the description of the conformation of polypeptide chains. Tentative rules (1969). *Biochemistry* 9: 3471–3479. PMID: [5509841](#)
5. Arnott S, Dover SD (1967) Refinement of bond angles of an alpha-helix. *J Mol Biol* 30: 209–212. PMID: [6077935](#)
6. Jeffrey GA, Saenger W (1991) Hydrogen bonding in biological structures. Berlin; New York: Springer-Verlag. xiv, 569 p. p.
7. Donohue J (1953) Hydrogen Bonded Helical Configurations of the Polypeptide Chain. *Proc Natl Acad Sci U S A* 39: 470–478. PMID: [16589292](#)
8. Miyazawa T (1961) Molecular vibrations and structure of high polymers. II. Helical parameters of infinite polymer chains as functions of bond lengths, bond angles, and internal rotation. *J Polymer Sci* 55: 215–231.
9. Nemethy G, Phillips DC, Leach SJ, Scheraga HA (1967) A Second Right-handed Helical Structure with the Parameters of the Pauling-Corey [alpha]-helix. *Nature* 214: 363–365. PMID: [6040611](#)
10. Baur WH (1965) On hydrogen bonds in crystalline hydrates. *Acta Crystallographica* 19: 909–916.
11. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, et al. (2000) The Protein Data Bank. *Nucleic Acids Res* 28: 235–242. PMID: [10592235](#)
12. Berman HM, Henrick K, Nakamura H (2003) Announcing the worldwide Protein Data Bank. *Nature Structural Biology* 10: 980. PMID: [14634627](#)
13. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67: 355–367. doi: [10.1107/S0907444911001314](#) PMID: [21460454](#)
14. Sheldrick GM, Schneider TR (1997) SHELXL: high-resolution refinement. *Methods Enzymol* 277: 319–343. PMID: [18488315](#)
15. Brunger AT (2007) Version 1.2 of the Crystallography and NMR system. *Nat Protoc* 2: 2728–2733. PMID: [18007608](#)
16. Nicholls RA, Long F, Murshudov GN (2012) Low-resolution refinement tools in REFMAC5. *Acta Crystallogr D Biol Crystallogr* 68: 404–417. doi: [10.1107/S090744491105606X](#) PMID: [22505260](#)
17. Wiberg KB, Marquez M, Castejon H (1994) Lone Pairs in Carbonyl Compounds and Ethers. *J Org Chem* 59: 6817–6822.
18. Ren PY, Ponder JW (2003) Polarizable atomic multipole water model for molecular mechanics simulation. *J Phys Chem B* 107: 5933–5947.
19. Stone AJ (1981) Distributed Multipole Analysis, or How to Describe a Molecular Charge-Distribution. *Chem Phys Lett* 83: 233–239.
20. Williams DE (1988) Representation of the molecular electrostatic potential by atomic multipole and bond dipole models. *J Comp Chem* 9: 745–763.

21. Williams DE (1991) Net Atomic Charge and Multipole Models for the ab initio Molecular Electric Potential. In: Boyd KBLaDB, editor. Rev Comput Chem. New York: VCH Publishers, Inc. pp. 219–271.
22. Stone AJ (2008) Intermolecular potentials. *Science* 321: 787–789. doi: [10.1126/science.1158006](https://doi.org/10.1126/science.1158006) PMID: [18687950](https://pubmed.ncbi.nlm.nih.gov/18687950/)
23. Ponder JW, Case DA (2003) Force fields for protein simulations. *Protein Simulations* 66: 27–85. PMID: [14631816](https://pubmed.ncbi.nlm.nih.gov/14631816/)
24. Best RB, Buchete NV, Hummer G (2008) Are current molecular dynamics force fields too helical? *Biophys J* 95: L7–L9.
25. Morozov AV, Kortemme T, Tsemekhman K, Baker D (2004) Close agreement between the orientation dependence of hydrogen bonds observed in protein structures and quantum mechanical calculations. *Proc Natl Acad Sci U S A* 101: 6946–6951. PMID: [15118103](https://pubmed.ncbi.nlm.nih.gov/15118103/)
26. Ponder JW, Wu C, Ren P, Pande VJ, Chodera JD, et al. (2010) Current Status of the AMOEBA Polarizable Force Field. *J Phys Chem* in press.
27. Ren P, Wu C, Ponder JW (2011) Polarizable Atomic Multipole-based Molecular Mechanics for Organic Molecules. *J Chem Theory Comput* 7: 3143–3161. PMID: [22022236](https://pubmed.ncbi.nlm.nih.gov/22022236/)
28. Schnieders MJ, Fenn TD, Pande VS, Brunger AT (2009) Polarizable atomic multipole X-ray refinement: application to peptide crystals. *Acta Crystallogr D Biol Crystallogr* 65: 952–965. doi: [10.1107/S0907444909022707](https://doi.org/10.1107/S0907444909022707) PMID: [19690373](https://pubmed.ncbi.nlm.nih.gov/19690373/)
29. Fenn TD, Schnieders MJ, Brunger AT, Pande VS (2010) Polarizable atomic multipole x-ray refinement: hydration geometry and application to macromolecules. *Biophys J* 98: 2984–2992. doi: [10.1016/j.bpj.2010.02.057](https://doi.org/10.1016/j.bpj.2010.02.057) PMID: [20550911](https://pubmed.ncbi.nlm.nih.gov/20550911/)
30. Zheng X, Wu C, Ponder JW, Marshall GR (2012) Molecular dynamics of beta-hairpin models of epigenetic recognition motifs. *J Am Chem Soc* 134: 15970–15978. PMID: [22934656](https://pubmed.ncbi.nlm.nih.gov/22934656/)
31. Ponder JW (2014) TINKER—Software Tools for Molecular Design. Available: <http://dasher.wustl.edu/tinker/>. Ponder Research Group, Department of Chemistry, Washington University in St. Louis.
32. Hollingsworth SA, Berkholz DS, Karplus PA (2009) On the occurrence of linear groups in proteins. *Protein Sci* 18: 1321–1325. doi: [10.1002/pro.133](https://doi.org/10.1002/pro.133) PMID: [19472372](https://pubmed.ncbi.nlm.nih.gov/19472372/)
33. Wlodawer A, Minor W, Dauter Z, Jaskolski M (2008) Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *Febs J* 275: 1–21. doi: [10.1111/j.1742-4658.2008.06444.x](https://doi.org/10.1111/j.1742-4658.2008.06444.x) PMID: [18564239](https://pubmed.ncbi.nlm.nih.gov/18564239/)
34. Blundell T, Barlow D, Borkakoti N, Thornton J (1983) Solvent-induced distortions and the curvature of alpha-helices. *Nature* 306: 281–283. PMID: [6646210](https://pubmed.ncbi.nlm.nih.gov/6646210/)
35. Perutz MF (1951) New x-ray evidence on the configuration of polypeptide chains. *Nature* 167: 1053–1054. PMID: [14843172](https://pubmed.ncbi.nlm.nih.gov/14843172/)
36. Arnott S, Wonacott AJ (1966) Atomic co-ordinates for an alpha-helix: refinement of the crystal structure of alpha-poly-L-alanine. *J Mol Biol* 21: 371–383. PMID: [5972467](https://pubmed.ncbi.nlm.nih.gov/5972467/)
37. Kuster DJ (2009) A Unifying Model for Helices in Proteins. St. Louis, MO: Washington University in St. Louis. 238 p.
38. Hovmoller S, Zhou T, Ohlson T (2002) Conformations of amino acids in proteins. *Acta Crystallogr D Biol Crystallogr* 58: 768–776. PMID: [11976487](https://pubmed.ncbi.nlm.nih.gov/11976487/)
39. Stockmayer WH (1967) Dielectric Dispersion in Solutions of Flexible Polymers. *Pure Appl Chem* 15. PMID: [5586552](https://pubmed.ncbi.nlm.nih.gov/5586552/)
40. Preissner R, Egner U, Saenger W (1991) Occurrence of bifurcated three-center hydrogen bonds in proteins. *FEBS Lett* 288: 192–196. PMID: [1879550](https://pubmed.ncbi.nlm.nih.gov/1879550/)
41. Stickle DF, Presta LG, Dill KA, Rose GD (1992) Hydrogen bonding in globular proteins. *J Mol Biol* 226: 1143–1159. PMID: [1518048](https://pubmed.ncbi.nlm.nih.gov/1518048/)
42. McDonald IK, Thornton JM (1994) Satisfying hydrogen bonding potential in proteins. *J Mol Biol* 238: 777–793. PMID: [8182748](https://pubmed.ncbi.nlm.nih.gov/8182748/)
43. Fain AV, Berezovsky IN, Chekhov VO, Ukrainkii DL, Esipova NG (2001) Double and Bifurcated Hydrogen Bonds in the  $\alpha$ -Helices of Globular Proteins. *Biophysics* 46: 921–928.
44. Fain AV, Ukrainkii DL, Dobkin SA, Galkin AV, Esipova NG (2008) Abundance and Arrangement of Single, Double and Bifurcated Hydrogen Bonds in Protein  $\alpha$ -Helices: Statistical Analysis of PDB-Select. *Biophysics* 53: 125–133.
45. Yuan Y, Mills MJ, Popelier PL (2014) Multipolar electrostatics for proteins: atom-atom electrostatic energies in crambin. *J Comput Chem* 35: 343–359. doi: [10.1002/jcc.23469](https://doi.org/10.1002/jcc.23469) PMID: [24449043](https://pubmed.ncbi.nlm.nih.gov/24449043/)
46. Cardamone S, Hughes TJ, Popelier PL (2014) Multipolar electrostatics. *Phys Chem Chem Phys* 16: 10367–10387. doi: [10.1039/c3cp54829e](https://doi.org/10.1039/c3cp54829e) PMID: [24741671](https://pubmed.ncbi.nlm.nih.gov/24741671/)

47. Kendrew JC (1959) Structure and function in myoglobin and other proteins. *Fed Proc* 18: 740–751. PMID: [13672267](#)
48. Perutz MF (1960) Structure of hemoglobin. *Brookhaven Symp Biol* 13: 165–183. PMID: [13734651](#)
49. Welsh E (2004) *A Set of Computational Tools for the Analysis and Prediction of Protein Structures*. St. Louis, MO: Washington University.
50. Fenn TD, Schnieders MJ, Mustyakimov M, Wu C, Langan P, et al. (2011) Reintroducing electrostatics into macromolecular crystallographic refinement: application to neutron crystallography and DNA hydration. *Structure* 19: 523–533. doi: [10.1016/j.str.2011.01.015](#) PMID: [21481775](#)
51. Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. *J Mol Biol* 7: 95–99. PMID: [13990617](#)
52. Quine JR (1999) Helix parameters and protein structure using quaternions. *J Mol Struct (Theochem)* 460: 53–66.
53. Tuckerman ME, Berne BJ, Rossi A (1991) Molecular dynamics algorithm for multiple time scales: Systems with disparate masses. *J Chem Phys* 94: 1465–1469.
54. Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* 314: 141–151.
55. Patriksson A, van der Spoel D (2008) A temperature predictor for parallel tempering simulations. *Phys Chem Chem Phys* 10: 2073–2077. doi: [10.1039/b716554d](#) PMID: [18688361](#)
56. Penev ES, Lampoudi S, Shea J-E (2009) TiReX: Replica-exchange molecular dynamics using Tinker. *Computer Physics Communications* 180: 2013–2019.
57. Teeter MM (1984) Water structure of a hydrophobic protein at atomic resolution: Pentagon rings of water molecules in crystals of crambin. *Proc Natl Acad Sci U S A* 81: 6014–6018. PMID: [16593516](#)
58. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79: 926–935.
59. Jelsch C, Teeter MM, Lamzin V, Pichon-Pesme V, Blessing RH, et al. (2000) Accurate protein crystallography at ultra-high resolution: Valence electron distribution in crambin. *Proc Natl Acad Sci U S A* 97: 3171–3176. PMID: [10737790](#)
60. Schmidt A, Teeter M, Weckert E, Lamzin VS (2011) Crystal structure of small protein crambin at 0.48 Å resolution. *Acta Crystallographica Section F* 67: 424–428. doi: [10.1107/S1744309110052607](#) PMID: [21505232](#)
61. Choudhary A, Raines RT (2011) Signature of  $n \rightarrow \pi^*$  interactions in alpha-helices. *Protein Sci* 20: 1077–1081. doi: [10.1002/pro.627](#) PMID: [21442680](#)
62. Liu C, Ponder JW, Marshall GR (2014) Helix Stability of Oligoglycine, Oligoalanine and Oligo- $\beta$ -alanine Dodecamers Reflected by Hydrogen-Bond Persistence. *Proteins* in press.
63. Shimizu K, Freitas AA, Farah JP, Dias LG (2005) Predicting hydration free energies of neutral compounds by a parametrization of the polarizable continuum model. *J Phys Chem A* 109: 11322–11327. PMID: [16331918](#)
64. Marshall GR, Hodgkin EE, Langs DA, Smith GD, Zabrocki J, et al. (1990) Factors governing helical preference of peptides containing multiple alpha, alpha-dialkyl amino acids. *Proc Natl Acad Sci U S A* 87: 487–491. PMID: [2296604](#)
65. Felder CE, Prilusky J, Silman I, Sussman JL (2007) A server and database for dipole moments of proteins. *Nucleic Acids Res* 35: W512–521. PMID: [17526523](#)
66. Sengupta D, Behera RN, Smith JC, Ullmann GM (2005) The alpha helix dipole: screened out? *Structure* 13: 849–855. PMID: [15939016](#)
67. Sitkoff D, Sharp KA, Honig B (1994) Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models. *J Phys Chem* 98: 1978–1988.
68. Ripoll DR, Vila JA, Scheraga HA (2005) On the orientation of the backbone dipoles in native folds. *Proc Natl Acad Sci U S A* 102: 7559–7564. PMID: [15894608](#)
69. Smythe ML, Huston SE, Marshall GR (1995) The Molten Helix—Effects of Solvation on the Alpha-Helical to 3(10)-Helical Transition. *J Am Chem Soc* 117: 5445–5452.
70. Smythe ML, Huston SE, Marshall GR (1993) Free-Energy Profile of a 3(10)-Helical to Alpha-Helical Transition of an Oligopeptide in Various Solvents. *J Am Chem Soc* 115: 11594–11595.
71. Bour P, Kubelka J, Keiderling TA (2002) Ab initio quantum mechanical models of peptide helices and their vibrational spectra. *Biopolymers* 65: 45–59. PMID: [12209472](#)
72. Pauling L, Corey RB (1951) The polypeptide-chain configuration in hemoglobin and other globular proteins. *Proc Natl Acad Sci U S A* 37: 282–285. PMID: [14834151](#)

73. Marshall GR, Bosshard HE (1972) Angiotensin II. Studies on the biologically active conformation. *Circ Res* 31: Suppl 2:143–150. PMID: [4341473](#)
74. Karle IL, Flippen-Anderson JL, Uma K, Balam P (1989) Modular design of synthetic protein mimics. Characterization of the helical conformation of a 13-residue peptide in crystals. *Biochemistry* 28: 6696–6701. PMID: [2790024](#)
75. Karle IL, Flippen-Anderson JL, Gurunath R, Balam P (1994) Facile transition between 3(10)- and alpha-helix: structures of 8-, 9-, and 10-residue peptides containing the-(Leu-Aib-Ala)<sub>2</sub>-Phe-Aib- fragment. *Protein Sci* 3: 1547–1555. PMID: [7833814](#)
76. Karle IL, Flippen-Anderson JL, Uma K, Balam H, Balam P (1989) Alpha-helix and mixed 3(10)/alpha-helix in cocrystallized conformers of Boc-Aib-Val-Aib-Aib-Val-Val-Val-Aib-Val-Aib-OMe. *Proc Natl Acad Sci U S A* 86: 765–769. PMID: [2915976](#)
77. Karle IL, Flippen-Anderson JL, Uma K, Balam P (1993) Unfolding of an alpha-helix in peptide crystals by solvation: conformational fragility in a heptapeptide. *Biopolymers* 33: 827–837. PMID: [8343578](#)
78. Aravinda S, Shamala N, Balam P (2008) Aib residues in peptaibiotics and synthetic sequences: analysis of nonhelical conformations. *Chem Biodivers* 5: 1238–1262. doi: [10.1002/cbdv.200890112](#) PMID: [18649312](#)
79. Paterson Y, Rumsey SM, Benedetti E, Nemethy G, Scheraga HA (1981) Sensitivity of polypeptide conformation to geometry. Theoretical conformational analysis of oligomers of alpha-aminoisobutyric acid. *Journal of the American Chemical Society* 103: 2947–2955.
80. Hodgkin EE, Clark JD, Miller KR, Marshall GR (1990) Conformational analysis and helical preferences of normal and alpha,alpha-dialkyl amino acids. *Biopolymers* 30: 533–546.
81. Shamala N, Nagaraj R, Balam P (1978) *J Chem Soc Chem Comm* 1978: 996–997.
82. Wieczorek R, Dannenberg JJ (2004) Comparison of fully optimized alpha- and 3(10)-helices with extended beta-strands. An ONIOM density functional theory study. *J Am Chem Soc* 126: 14198–14205. PMID: [15506786](#)
83. Parra RD, Zeng H, Zhu J, Zheng C, Zeng XC, et al. (2001) Stable three-center hydrogen bonding in a partially rigidified structure. *Chemistry* 7: 4352–4357. PMID: [11695668](#)
84. Rozas I, Alkorta I, Elguero J (1998) Bifurcated Hydrogen Bonds: Three-Centered Interactions. *J Phys Chem A* 102: 9925–9932.
85. Yang J, Christianson LA, Gellman SH (1999) Comparison of an HXH Three-Center Hydrogen Bond with Alternative Two-Center Hydrogen Bonds in a Model System. *Org Lett* 1: 11–13.
86. Feldblum ES, Arkin IT (2014) Strength of a bifurcated H bond. *Proceedings of the National Academy of Sciences* 111: 4085–4090. doi: [10.1073/pnas.1319827111](#) PMID: [24591597](#)
87. Stone AJ (2005) Distributed multipole analysis: Stability for large basis sets. *J Chem Theory Comput* 1: 1128–1132.
88. Hol WG (1985) The role of the alpha-helix dipole in protein function and structure. *Prog Biophys Mol Biol* 45: 149–195. PMID: [3892583](#)
89. Scheraga HA (1985) Effect of side chain-backbone electrostatic interactions on the stability of alpha-helices. *Proc Natl Acad Sci U S A* 82: 5585–5587. PMID: [3862082](#)
90. Shoemaker KR, Kim PS, Brems DN, Marqusee S, York EJ, et al. (1985) Nature of the charged-group effect on the stability of the C-peptide helix. *Proc Natl Acad Sci U S A* 82: 2349–2353. PMID: [3857585](#)
91. Panman MR, Bakker BH, den Uyl D, Kay ER, Leigh DA, et al. (2013) Water lubricates hydrogen-bonded molecular machines. *Nat Chem* 5: 929–934. doi: [10.1038/nchem.1744](#) PMID: [24153370](#)
92. Hunt NT, Kattner L, Shanks RP, Wynne K (2007) The Dynamics of Water-Protein Interaction Studied by Ultrafast Optical Kerr-Effect Spectroscopy. *Journal of the American Chemical Society* 129: 3168–3172. PMID: [17315992](#)