AlphaFold2 captures the conformational landscape of the HAMP signaling domain

Aleksander Winski[a], Jan Ludwiczak[b,†], Malgorzata Orlowska[b], Rafal Madaj[b], Kamil Kaminski[b], and Stanislaw Dunin-Horkawicz[b,c,*]

[a] Laboratory of Structural Bioinformatics, Centre of New Technologies, University of Warsaw, Banacha 2c, 02-097 Warsaw, Poland

[b] Institute of Evolutionary Biology, Faculty of Biology, Biological and Chemical Research Centre, University of Warsaw, Żwirki i Wigury 101, 02-089, Warsaw, Poland

[c] Department of Protein Evolution, Max Planck Institute for Biology Tübingen, Max-Planck-Ring 5, 72076 Tübingen, Germany

[†] Present address: Prescient Design, Genentech Research & Early Development, Roche Group, Grenzacherstrasse 124, 4070 Basel, Switzerland.

[*] To whom correspondence should be addressed
E-mail: stanislaw.dunin-horkawicz@tuebingen.mpg.de

Abstract

Computational modeling of protein structures that adopt multiple conformations is a challenging task. We propose a procedure for generalizing the knowledge gained from AlphaFold2 models to represent such conformational landscapes. The procedure was applied to the HAMP domain, a short helical bundle that transduces the signal from sensors to effectors in two-component signaling proteins such as EnvZ or Tar. The results obtained correctly reproduce the experimentally confirmed conformations and support the universality of the mechanism by which HAMP domains transduce signal by axial rotation of their helices. Furthermore, the AlphaFold2 models were used to train a deep learning tool, HAMPpred, which can predict the resting conformation of a given HAMP domain within seconds. Using HAMPpred, we searched 60,000 HAMP domains for cases where two HAMP sequences differing by one amino acid showed a significant change in predicted conformation. We found that these likely adaptive mutations occur in proteins involved in processes such as antibiotic production and cell wall damage sensing, or in proteins containing rapidly evolving poly-HAMP arrays. All code associated with this work is deposited at https://github.com/labstructbioinf/HAMPpred.

Introduction

The HAMP domain is an approximately 50-amino acid protein domain consisting of two helices and an unstructured region connecting them (Dunin-Horkawicz and Lupas, 2010a; Hulko *et al.*, 2006). It is mostly found in dimeric signal transduction proteins, such as prokaryotic sensor histidine kinases and chemoreceptors, where it links the sensory extracellular domain to the downstream effector and acts as a signal transduction modulator (Figure 1). In the context of a dimer, the HAMP domain forms a four-helical coiled coil (Lupas *et al.*, 2017) in which all helices are oriented in parallel. The coiled-coil domains are characterized by the presence of the seven-residue (heptad) sequence repeat $[abcdefg]_n$, in which residues at positions *a* and *d* face the center of the bundle. These residues from opposite helices interact via the knobs-in-holes packing (Lupas *et al.*, 2017) to form zipper-like structures.

The HAMP domain has been extensively studied both experimentally and computationally, leading to models that describe the conformational changes it undergoes during signaling (Ferris *et al.*, 2012; Hulko *et al.*, 2006; Parkinson, 2010; Airola *et al.*, 2010; Matamouros *et al.*, 2015). One such model describes these changes as a coordinated rotation of helices between two boundaries (Hulko *et al.*, 2006). For the well-studied HAMP domain of the *Archaeoglobus fulgidus* Af1503 protein, the rotation boundaries are defined by two states: the canonical knobs-into-holes packing, in which two residues form the hydrophobic core (Figure 2B) (Ferris *et al.*, 2012), and an alternative form of packing, the x-da packing, in which an additional residue is co-opted into the core (Figure 2A) (Hulko *et al.*, 2006). The transition from the canonical to the x-da packing is achieved by rotating the N- and C-terminal helices counterclockwise and clockwise, respectively, relative to the knobs-in-holes packing (compare Figure 2A and B). These changes can be quantified (Szczepaniak *et al.*, 2021; Dunin-Horkawicz and Lupas, 2010b) and described as the relative rotation between the N- and C-helices. The rotation value for the canonical packing is 0°, while the x-da packing has a value of -26°.

Sequence analyses suggest that the signaling model proposed for Af1503 may apply to most HAMP domains (Dunin-Horkawicz and Lupas, 2010a). However, some HAMP domains are postulated to adopt conformations between the canonical packing (as in the case of Af1503) and an alternative packing, da-x (rotation of +26°), in which the helices are rotated in the opposite direction (N- and C-terminal helices are rotated clockwise and counterclockwise, respectively, relative to the knobs-in-holes packing; Figure 2C). Such alternative HAMP domains are often found in association with the canonical ones, forming poly-HAMP arrays in which the two forms alternate (Airola *et al.*, 2010; Anaya *et al.*, 2022). This alternating pattern may be due to the mechanism of HAMP function, in which the N- and C-terminal helices rotate in opposite directions when switching between the two conformations (Airola *et al.*, 2013; Dunin-Horkawicz and Lupas, 2010a). Therefore, HAMP domains of the same type cannot function together because this would require one of them to adopt an inaccessible conformation.

The conformations described above, namely the canonical knobs-in-holes, x-da and da-x packings, correspond to idealized rotational states of 0°, -26°, and +26°, respectively. However, HAMP domains can adopt a continuum of conformations (Dunin-Horkawicz and Lupas, 2010b), ranging from x-da to knobs-into-holes in canonical HAMPs and from knobs-into-holes to da-x in the alternative HAMPs. To explore this conformational landscape, we used AlphaFold2 to model >5000 HAMP domains from 81 families (Dunin-Horkawicz and Lupas, 2010a) and investigated how these models can contribute to our understanding of HAMP domain structure and function.

Materials and Methods

AlphaFold2 modelling

6456 aligned HAMP sequences were obtained from (Dunin-Horkawicz and Lupas, 2010a), and all that contained gaps in helices were removed. Sequences were then clustered at 70% identity to remove redundancy, resulting in a final set of 5388 sequences. The HAMP domains of this set were modeled as dimers using the ColabFold 1.5.2 environment (Mirdita *et al.*, 2022) with default parameters (no templates were used), and the rotation of their helices was measured using SamCC Turbo (Figure 2) (Szczepaniak *et al.*, 2021). This tool

analyzes a given coiled-coil structure by dividing it into layers comprising *n* Cα atoms (where *n* is the number of helices in a bundle) lying approximately in a plane perpendicular to the bundle axis. The structure parameters are then calculated separately for each layer. The helix axial rotation, which is essential for describing the HAMP domain conformation, is calculated based on the Crick angles, which quantify the orientation of residues with respect to the bundle axis (0° indicates a residue pointing directly toward the bundle axis, while the +/-180° indicates a residue pointing directly outward from the bundle axis). The helix rotation is calculated as the difference between the measured Crick angles and those expected for a canonical coiled coil assuming knobs-in-holes packing. The expected values were defined individually for each heptad position as a:19.5°, b:122.35°, c:-134.78°, d:-31.92°, e:70.92°, f:173.78°, and g:-83.35°.

In the case of HAMP domain, the N- and C-terminal helices are considered separately because they rotate in the opposite directions (Figure 2). The Crick angles were measured for the 11 residues of N- and C-terminal HAMP helices and compared to the reference angles to quantify their rotation (the match between the reference heptad register and the heptad positions of the modelled helices was defined based on the multiple sequence alignment). In this way, each HAMP model was assigned with two 11-element vectors describing the rotation of N and C-terminal helices. Values close to zero indicate rotation compatible with knobs-into-holes packing, while values above or below zero indicate da-x or x-da packing(Figure 2). Finally, global helix axial rotation was defined as the difference between the rotation of the N- and C-terminal helices divided by two (see black solid lines in Figure 2). The 11-element global helix axial rotation vector represents the relative rotation of the N/C-terminal helices and is independent of the periodicity of the bundle. Finally, by averaging this vector, a HAMP structure conformation can be represented by a single number ranging from -26° to +26° (Figure 3B and C).

HAMPpred prediction model

To enable the rapid prediction of HAMP domain conformation (i.e., global helix axial rotation) based on sequence, a deep learning model was developed. AlphaFold2 HAMP models with an absolute mean global rotation greater than 26 or a standard deviation of global rotation greater than 14 were removed, resulting in a set of 4972 models. In each

model, the sequences of the N- and C-terminal helices (model input) were encoded separately as a matrix of shape (11, 23), where 11 is the helix length and 23 is the one-hot encoded amino acid and its hydrophobicity and side-chain size (we used the Kyte and Doolittle hydrophobicity scale (Kyte and Doolittle, 1982) and obtained side-chain sizes from the Jena Library of Biological Macromolecules (Huehne and Suehnel, 2009)). The encodings for the N- and C-terminal helices were concatenated along the last dimension, resulting in a matrix of shape (11, 46) for each model. The global helix axial rotation vectors (model output) were encoded using sin and cos functions, resulting in a matrix of the form (11, 2), where 11 is the global helix axial rotation vector length and 2 is the sin and cos of its values. Such encoding avoids periodic effects of degrees and can be easily inverted using the arc tangent function.

This set of 4972 encoded HAMP sequences (input) and conformations (output) was randomly divided into 3978 training and 994 validation examples. The training set was used to train 4 model variants, while the validation set was used for early termination and selection of the best model architecture. The general model architecture was 3 convolutional filters, each consisting of 3 parallel layers with kernel sizes of 3, 4, and 7 and a batch normalization layer, 1 or 2 bidirectional LSTM layers, 1 or 3 dense layers, and an output layer. Each model was trained for 100 epochs with an initial learning rate of 0.0001. Tanh was used as the activation function and mean square error (MSE) as the loss function. If the validation loss did not improve for 7 epochs, the learning rate was set to a value 10 times lower than the previous one. If the validation loss did not improve for 15 epochs, training was stopped. The best model had two LSTM layers and a single dense layer.

Classification of di-HAMP modules

The GAF-DHp-HisKA domain segments were identified in the poly-HAMP proteins A0A1N6G1D1, K9PBI3, G0G0R7, A0A4R4P399, D8FVA6, A0A0X3XHH5, A0A248K2C0, A0A1H5PQZ5, A0A2W2M3F7, I4YVW6, A0A2U1VYS9, A0A542ZIX7, A0A1U7JAJ8, B0UB08, A0A4Z0BGY9, A0A2D3U389, A0A1Q9RWB5, A0A1G9HBX6, A0A3E0HHV0, A0A0X3RTC7, A0A1N7CPU7, A0A2A3HZJ6, and A0A3A9YLK8. These were then used to search the NCBI NR database using BLAST (Altschul *et al.*, 1997) with a cut-off e-value of 1e-5. Only unique hits with query coverage of at least 90% and sequence identity of at least 60% were considered

for further analysis. The full-length protein sequences of the obtained hits were used to construct a custom BLAST database. Next, we manually identified the individual di-HAMPs (canonical plus alternative HAMP modules) in the K9PBI3 and A0A1N6G1D1 sequences and used them as queries to search the custom database with BLAST (e-value cut-off 1e-5, query coverage of at least 90%, sequence identity of at least 40%). We discarded duplicates and partial hits, and the resulting di-HAMPs were aligned using Clustal Omega (Sievers *et al.*, 2011). From the resulting multiple sequence alignment, we extracted first (canonical) HAMPs from each di-HAMP and clustered them using CLANS (Frickey and Lupas, 2004) (Figure 6).

Characterization of HAMP domain families

HAMP domain families were defined based on the previous classification (Dunin-Horkawicz and Lupas, 2010a). To assess the conformational preference of HAMP domains from a given family, a normalized distribution of rotation (averaged global helix axial rotation) across all models was calculated and then subtracted from the distribution calculated for the family of interest (Figure 4). Such a comparison allows highlighting rotational states that are more (enrichment) or less (depletion) common in a given family compared to the background of all HAMP models.

Results and discussion

We modeled 11 HAMP domain dimers (Table 1) using AlphaFold2 (Mirdita *et al.*, 2022) and compared the resulting models with corresponding experimental structures. The structure comparisons were performed using the SamCC tool (Szczepaniak *et al.*, 2021), which allows the measurement of helical rotation in coiled coils (see Methods). The results obtained are consistent with previous observations (Anaya *et al.*, 2022) and indicate a high accuracy of AlphaFold2, which models show conformations very similar to those observed in the experimental structures (Figure 3). The only outlier was a variant of *Archaeoglobus fulgidus* Af1503 HAMP in which the alanine at position 291 was replaced by phenylalanine to force the knobs-into-holes canonical conformation. The difficulty in modeling this case is probably related to the fact that the phenylalanine at this position is never observed in the Af1503

homologs (in fact, its introduction may even cause the Af1503 HAMP to switch to the non-functional anti-parallel conformation (Ferris *et al.*, 2011)).

Given the good performance of AlphaFold2 on a small set (Table 1, Figure 3), we calculated models of >5000 HAMP domains representing the 81 families identified in our previous work (Dunin-Horkawicz and Lupas, 2010). We found that in most families, HAMP domains show a conserved preference for a resting conformation. For example, 200 HAMP domains (average sequence identity of 35%) grouped in the "B" family, found almost exclusively in chemoreceptors, are biased toward *x-da* packing (see green line in Figure 4A). We hypothesize that this conserved preference reflects the "tuning" of a given HAMP family to function in the particular context of other domains, and that these domains co-evolve to maintain a balance between their conformational preferences.

Such co-evolution can be clearly seen in the case of di-HAMP and poly-HAMP modules, i.e. two or more HAMP domains chained together. For example, some of the "B" family HAMPs are followed by a HAMP from another family to form a di-HAMP module. According to the rotation model (Hulko *et al.*, 2006), the two domains of a di-HAMP module are expected to rotate in opposite directions and thus should have opposite resting state preferences. Indeed, an opposite preference is observed for F2-HAMPs, which are often found downstream of B-HAMPs (compare green and orange lines in Figure 4A). In such B-F2 modules, the B-HAMPs would rotate between x-da and knobs-into-holes conformations, whereas the F2-HAMPs would rotate between da-x and knobs-into-holes conformations (Figure 2).

While B-type HAMPs are found both alone and in association with F2-type HAMPs, the G1 and F6 HAMP domains almost exclusively form long poly-HAMP arrays in which sequences from both groups alternate. The results obtained (Figure 4B) indicate that HAMP domains from these two groups have clearly opposite preferences for resting state conformation, with the canonical ones (F6 group) being shifted towards x-da packing and the alternative ones (G1 group) towards da-x packing. We propose that the conserved conformational preferences of the HAMP families could be used to aid in the design of hybrid proteins in which a HAMP domain is placed in a new context (e.g., upstream of another output domain

or in conjunction with other HAMPs). We also envision that analogous compatibility rules could be derived for other coiled-coil domains of two-component signaling systems.

As described above, the AlphaFold2 models provide insight into the function and evolution of HAMP domains; however, their applicability can be limited by the time required to run folding simulations. This is particularly true for high-throughput scans involving thousands of natural or mutant sequences. To address this problem, we used HAMP AlphaFold2 structures to train a prediction model, HAMPpred, which is capable of predicting the resting conformation from sequences alone within seconds. The benchmark using 11 experimental structures (Table 1) showed that HAMPpred provides comparable accuracy to and AlphaFold2, with the main difference being that HAMPpred more accurately predicts the effect of the A291F mutation (Figure 3C).

There is considerable experimental evidence that the introduction of point mutations in HAMP domains can significantly affect the functionality of the entire receptor in which they occur (Matamouros *et al.*, 2015; Hulko *et al.*, 2006; Mondéjar *et al.*, 2012; Ferris *et al.*, 2012). We hypothesized that such changes may also occur naturally as a result of adaptive evolution. To address this question, we used HAMPpred to analyze 1500 pairs of HAMP sequences from the Pfam family PF00672 that differ by only one amino acid at one of the hydrophobic core positions (Figure 2). For all such pairs, we calculated the absolute difference in their predicted rotational states (Figure 5A). For most (>90%) pairs, the predicted difference was less than 2°, consistent with the observation that the resting conformation is conserved in HAMP families. However, in some cases the difference was larger, up to 7°. We examined those with differences in the top 3% and found that some of them could be associated with functions that may be under selective pressure. For example, we identified a pair of sensor kinases from *Streptomyces niveus* and *Streptomyces agglomeratus* that are putative homologs of CseC, a sensor kinase involved in the response to cell envelope damage in *Streptomyces coelicolor* (Hong *et al.*, 2002) (Figure 5B). Another example also comes from the genus Streptomyces, where we identified a pair of kinases from *Streptomyces* sp. TSRI0281 and *Streptomyces* sp. Z26 that are homologs of MtrB, a receptor kinase involved in the regulation of antibiotic production in *Streptomyces coelicolor*

and *Streptomyces venezuelae* (Som, Heine, N. A. Holmes, *et al.*, 2017; Som, Heine, N. Holmes, *et al.*, 2017) (Figure 5B).

The pronounced effects of point mutations were also observed in the canonical HAMP domains that are part of the poly-HAMP arrays (the alternative HAMP domains were not analyzed because they are not covered by the Pfam PF00672 family). In the poly-HAMP arrays, the initial and terminal HAMP domains are distinct from each other and from those in between (Figure 6). In contrast, the middle HAMP domains are very similar to each other, suggesting that they have recently evolved by amplification. For example, the middle HAMP domains of the poly-HAMP array of a putative receptor from *Methylobacterium* sp. (B0UB08) share more than 90% sequence identity on average. The presence of such putative adaptive mutations exclusively in the middle HAMP domains may be related to the recovery of conformational equilibrium after amplification.

We also found that the predicted adaptive mutations often occur in different but very similar sequence pairs at different positions, suggesting that these proteins are generally susceptible to such changes. In cases such as CseC and MtrB homologs, this may be related to their vital function, whereas in the case of poly-HAMP arrays, it may be related to the need to adapt to the effects of amplification.

In summary, we have shown that AlphaFold2 is a powerful tool for analyzing protein structures that adopt different conformations. AlphaFold2 not only reproduces the conformations observed in experimental structures, but also allows to map the conformational landscape of the entire protein family. The AlphaFold2-based procedure was applied to the HAMP domain family and yielded results consistent with the rotational model of signal transduction. We believe that the presented approach can be applied to any protein family whose conformations can be quantified.

Acknowledgments

This work was supported by „Diamentowy Grant" DI2018010748 (to A.W.) from the Polish Ministry of Science and Higher Education. S.D-H. was supported by institutional funds from

the Max Planck Society. The authors would like to thank Andrei Lupas and Murray Coles for their constructive comments and critical evaluation of the manuscript.

Tables

Table 1. Experimental HAMP structures showing different types of hydrophobic core packing. Numbers indicate start and end of N- and C-terminal helices in PDB numbering. Letters above the sequences describe the heptad positions for the canonical knobs-into-holes packing (see Figure 2B).

PBD ID	N start	N sequence	N end	C start	C sequence	C end
		a d a d			*a d a d*	
2l7h	284	IIELSNTADKI	294	312	IGILAKSIERL	322
2l7i	284	IIELSNTFDKI	294	312	IGILAKSIERL	322
2y0q	284	IIELSNTCDKI	294	312	IGILAKSIERL	322
2y20	284	IIELSNTIDKI	294	312	IGILAKSIERL	322
2y21	284	IIELSNTVDKI	294	312	IGILAKSIERL	322
5iji	180	LNQLVTASQRI	190	208	LGLLAKTFNQM	218
4biv	180	ARKLKNAADEV	200	218	FLAAGASFNQM	228
4i5s	41	VKQLNAKVRSL	51	68	LSELVNNVNDL	78
3lnr	14	ADRIATLLQSF	24	41	YERLYDSLRAL	51
3lnr	69	EAGLAEMSRQH	79	97	AARIAKGVNEL	107
3lnr	115	KMKVVSVVTAY	125	141	KAQITEAIDGV	151

Figures

Figure 1. AlphaFold2 models of the E. coli Tar and EnvZ dimeric receptors. In both models, the HAMP domain is shown in color, with its N- and C-terminal helices shown in red and cyan, respectively.

Figure 2. The three basic conformations observed in the experimental HAMP structures. (A) The WT Af1503 HAMP domain adopts an x-da packing in which three residues form the hydrophobic core. (B) The A291F variant of the Af1503 HAMP domain adopts the canonical coiled-coil packing in which two residues form the hydrophobic core. (C) The second HAMP of the Aer2 protein poly-HAMP array adopts a da-x packing in which three residues form the hydrophobic core. The left panels show cross sections of HAMP domain structures viewed

from the N-terminus (linkers between N- and C-terminal helices are not shown), with letters indicating heptad positions. The middle panels show the corresponding wheel diagrams showing which heptad positions are involved in the formation of the hydrophobic core. The right panels show measurements of the axial rotation of the N- (red) and C-terminal (cyan) helices. The black lines indicate the global rotation, calculated as the difference between the rotation of the N- and C-terminal helices.

Figure 3. Performance of AlpahFold2 in modeling HAMP domains. (A) Modeling of the four HAMP domains (three of which are also shown in Figure 2). The black lines show the global rotation measured in the experimental structures, while the light green lines show the global rotation measured in the AlphaFold2 models. (B) Modeling of the 11 HAMP domains listed in Table 1. The X and Y axes show the averaged global rotation in the AlphaFold2 models and the experimental structures, respectively. (C) Sequence-based prediction of the conformation of the HAMP domains. The two panels correspond to two prediction models. The x and y axes show the averaged global rotation predicted by HAMPpred and the averaged global rotation measured in the experimental structures, respectively.

Figure 4. Conformational preferences in HAMP domain families. The plots show the difference between the rotation distributions measured in all 5388 HAMP models and in a single HAMP family. A deviation from the baseline indicates that a given HAMP family is depleted or enriched in certain conformational states. (A) HAMP type B domains occur alone or in combination with HAMP type F2 domains to form di-HAMP modules. (B) G1 and F6 HAMPs exclusively form poly-HAMP arrays in which the G1-F6 module can be repeated several times.

Figure 5. Predicted effects of point mutations in the hydrophobic core of HAMP domains. (A) Each point represents a pair of HAMP domains that differ by only one residue. The position of each point along the y-axis indicates the absolute difference between their predicted rotational states. The red horizontal lines represent the top 3% of these differences. (B) Examples of the top 3% of HAMP domain pairs. Mutations are shown in red.

Figure 6. Cluster map of canonical HAMP domains found in poly-HAMP arrays. Each dot represents a single canonical HAMP domain sequence (the alternative HAMPs were not included). The relative position of the dots reflects sequence similarity as measured by BLAST. Colored dots indicate individual canonical HAMP domains of a poly-HAMP array shown below, and the one in which a potential adaptive mutation was detected is shown in red. Note that the first and last two HAMP domains are from separate clusters, while those in between are in the central cluster.

Bibliography

Airola,M. V. *et al.* (2013) HAMP Domain Conformers That Propagate Opposite Signals in Bacterial Chemoreceptors. *PLoS Biol.*, **11**, e1001479.

Airola,M. V. *et al.* (2010) Structure of Concatenated HAMP Domains Provides a Mechanism for Signal Transduction. *Structure*, **18**, 436–448.

Altschul,S.F. *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–402.

Anaya,S. *et al.* (2022) New Roles for HAMP Domains: the Tri-HAMP Region of Pseudomonas aeruginosa Aer2 Controls Receptor Signaling and Cellular Localization. *J. Bacteriol.*, **204**, e0022522.

Dunin-Horkawicz,S. and Lupas,A.N. (2010a) Comprehensive analysis of HAMP domains: implications for transmembrane signal transduction. *J. Mol. Biol.*, **397**, 1156–74.

Dunin-Horkawicz,S. and Lupas,A.N. (2010b) Measuring the conformational space of square four-helical bundles with the program samCC. *J. Struct. Biol.*, **170**, 226–35.

Ferris,H.U. *et al.* (2012) Mechanism of regulation of receptor histidine kinases. *Structure*, **20**, 56–66.

Ferris,H.U. *et al.* (2011) The mechanisms of HAMP-mediated signaling in transmembrane receptors. *Structure*, **19**, 378–85.

Frickey,T. and Lupas,A. (2004) CLANS: a Java application for visualizing protein families based on pairwise similarity. *Bioinformatics*, **20**, 3702–4.

Hong,H.-J. *et al.* (2002) A signal transduction system in Streptomyces coelicolor that activates the expression of a putative cell wall glycan operon in response to vancomycin and other cell wall-specific antibiotics. *Mol. Microbiol.*, **44**, 1199–1211.

Huehne,R. and Suehnel,J. (2009) The Jena Library of Biological Macromolecules - JenaLib. *Nat. Preced.*

Hulko,M. *et al.* (2006) The HAMP domain structure implies helix rotation in transmembrane signaling. *Cell*, **126**, 929–40.

Kyte,J. and Doolittle,R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.*, **157**, 105–32.

Lupas,A. *et al.* (2017) The Structure and Topology of α-Helical Coiled Coils. In, *Sub-Cellular Biochemistry*., pp. 95–129.

Matamouros,S. *et al.* (2015) HAMP Domain Rotation and Tilting Movements Associated with Signal Transduction in the PhoQ Sensor Kinase. *MBio*, **6**, e00616-15.

Mirdita,M. *et al.* (2022) ColabFold: making protein folding accessible to all. *Nat. Methods*, **19**, 679–682.

Mondéjar,L.G. *et al.* (2012) HAMP domain-mediated signal transduction probed with a mycobacterial adenylyl cyclase as a reporter. *J. Biol. Chem.*, **287**, 1022–31.

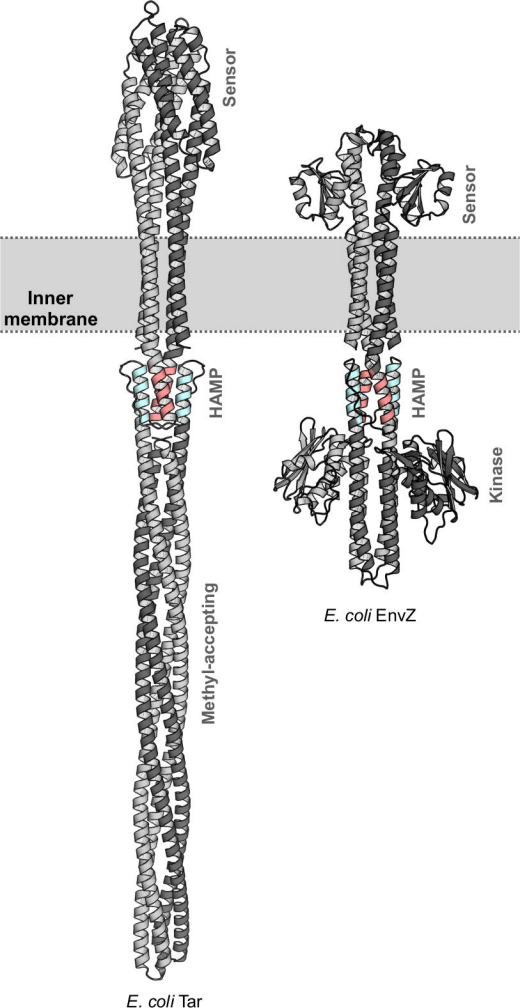
Parkinson,J.S. (2010) Signaling mechanisms of HAMP domains in chemoreceptors and sensor kinases. *Annu. Rev. Microbiol.*, **64**, 101–22.

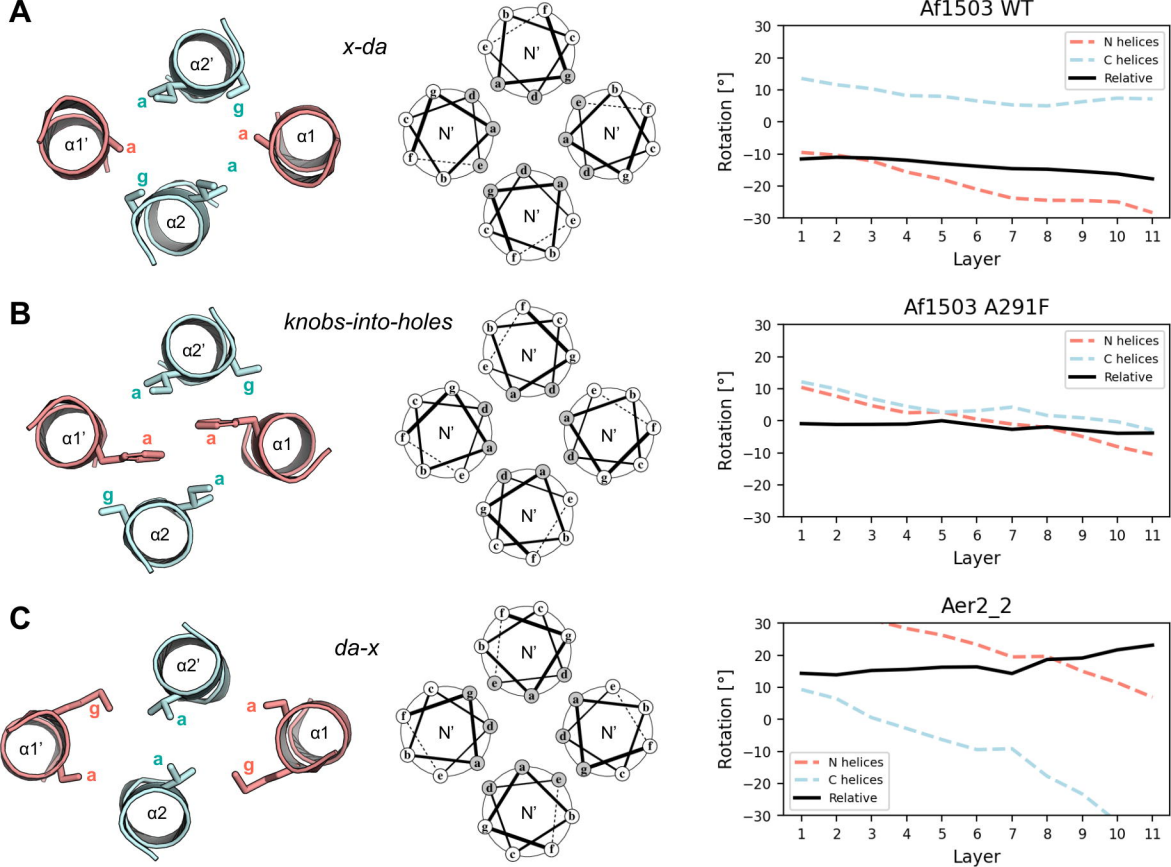
Sievers,F. *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.*, **7**, 539.

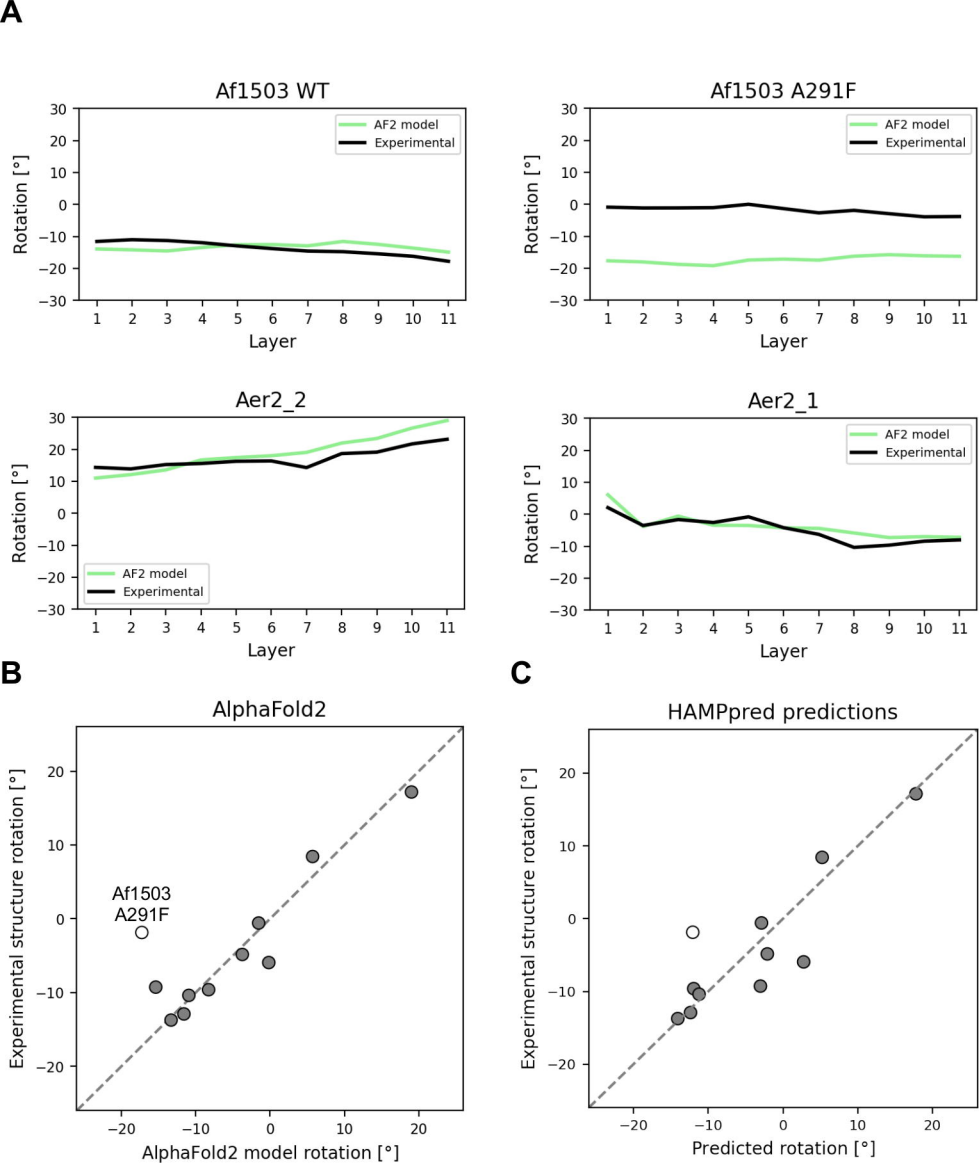
Som,N.F., Heine,D., Holmes,N.A., *et al.* (2017) The Conserved Actinobacterial Two-Component System MtrAB Coordinates Chloramphenicol Production with Sporulation in Streptomyces venezuelae NRRL B-65442. *Front. Microbiol.*, **8**, 1145.

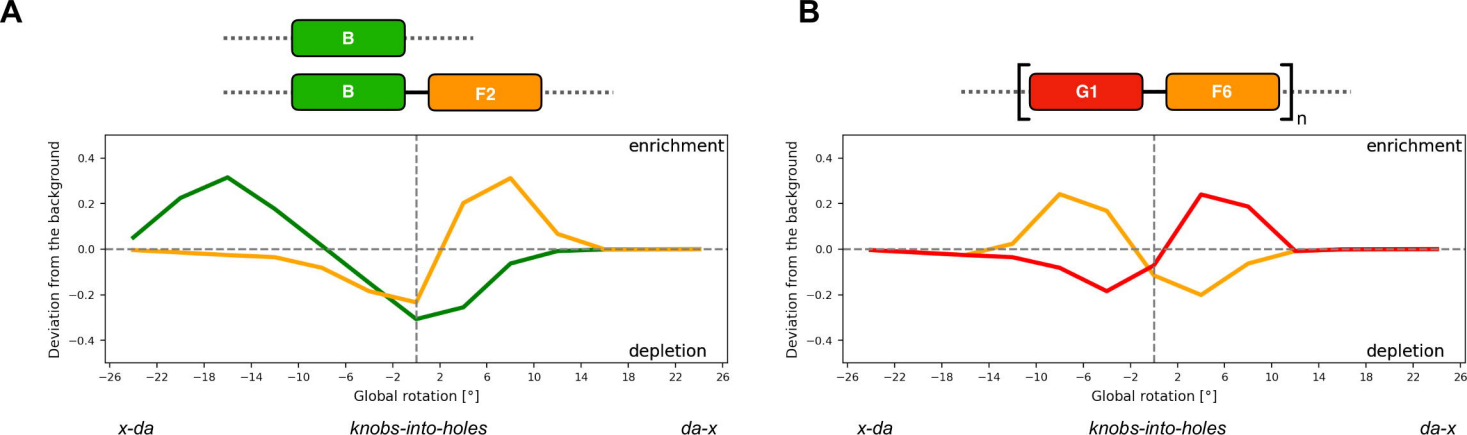
Som,N.F., Heine,D., Holmes,N., *et al.* (2017) The MtrAB two-component system controls antibiotic production in Streptomyces coelicolor A3(2). *Microbiology*, **163**, 1415–1419.

Szczepaniak,K. *et al.* (2021) A library of coiled-coil domains: from regular bundles to peculiar twists. *Bioinformatics*, **36**, 5368–5376.

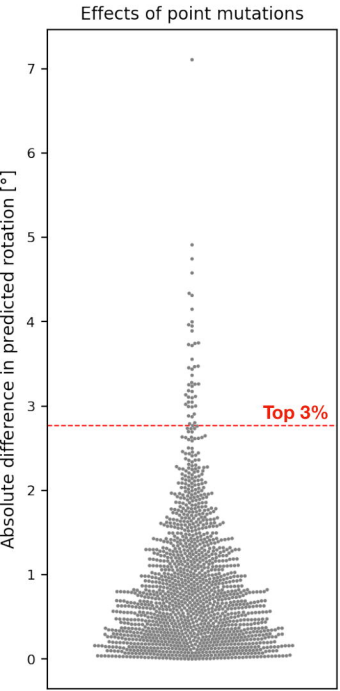

Sensor
Inner membrane
HAMP
Methyl-accepting
E. coli Tar
Sensor
HAMP
Kinase
E. coli EnvZ


A
x-da
α2'
α1'
α1
α2
a
g
N'
Af1503 WT
N helices
C helices
Relative
Rotation [°]
Layer
B
knobs-into-holes
Af1503 A291F
C
da-x
Aer2_2


A
Af1503 WT
Af1503 A291F
Aer2_2
Aer2_1
AF2 model
Experimental
Rotation [°]
Layer
B
AlphaFold2
Af1503
A291F
Experimental structure rotation [°]
AlphaFold2 model rotation [°]
C
HAMPpred predictions
Experimental structure rotation [°]
Predicted rotation [°]


A
B
B
F2
enrichment
depletion
Deviation from the background
0.4
0.2
0.0
−0.2
−0.4
−26
−22
−18
−14
−10
−6
−2
2
6
10
14
18
22
26
Global rotation [°]
x-da
knobs-into-holes
da-x
B
G1
F6
n
enrichment
depletion
Deviation from the background
0.4
0.2
0.0
−0.2
−0.4
−26
−22
−18
−14
−10
−6
−2
2
6
10
14
18
22
26
Global rotation [°]
x-da
knobs-into-holes
da-x

A



B

cseC — 4.7°

```
a  d   a  d   a  d   a  d
LRKAAAAAGKV...TDELARAVDAL V6KEP3 Streptomyces niveus NCIMB 11891
||||||||||||   ||||||||||
LRKAAAAAGKV...TDELARAVDAM A0A1E5P8C5 Streptomyces agglomeratus
```

mtrB — 4.3°

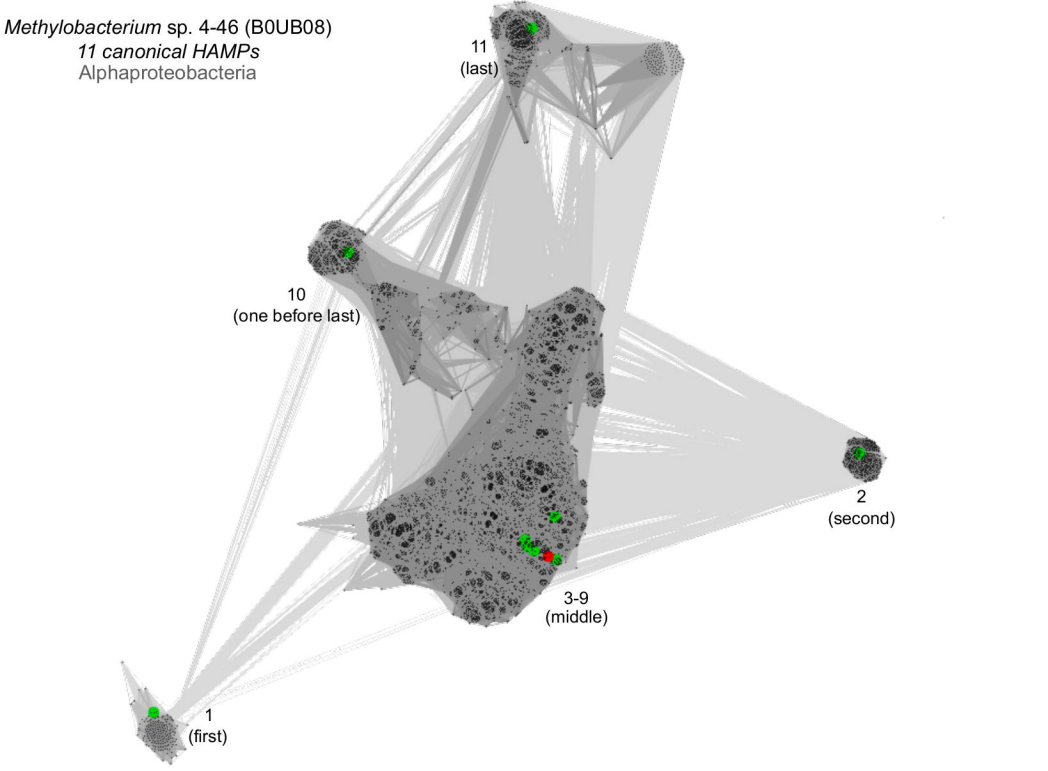
```
a  d   a  d   a  d   a  d
VRMAAGIVERL...IARLGEAFNKM A0A1Q4YVI0 Streptomyces sp. TSRI0281
||||||| |||   ||||||||||||
VRMAAGISERL...IARLGEAFNKM A0A498DN17 Streptomyces sp. Z26
```

poly-HAMP — 3.7°

```
a  d   a  d   a  d   a  d
VRNIATVTTAV...ILELKETINTM A0A519EQI2 Variovorax sp.
||||||||||||   ||||||| |||
VRNIATVTTAV...ILELKETFNTM A0A4Z0BGY9 Ramlibacter rhizophilus
```

poly-HAMP — 3°

```
a  d   a  d   a  d   a  d
VRNIAEVTTAV...ILDLKSTINTM A0A1U7JAJ8 Phormidium tenue
||||||||||||   ||||||| |||
VRNIAEVTTAV...ILDLKSTFNTM B0UB08 Methylobacterium sp. 4-46
```




```
      a  de  a  de                    ga  d  ga  d      ga  d  ga  d                  a  de  a  de
ATAQLEPKLLLKTLRAFRKGDFSVRLPLDLT-----GIDGEIAEALNDVVELNQSLARELDRVARKVGKDGRIAERGKLPTATGGWVECIDSVNAMI
GDLVQPTTEVARVIGAVAKGDLGQTMQLEIEGRPLQGEFLRIGKVVNTMVDQLNSFASEVTRVAREVGSEGKLGGQAQVKGVGGTWKDLTENVNLMA
ANLTGQVRNIAEVTTAVANGDLSKKITVDVK-----GEILDLKSTFNTMVDQLNSFASEVTRVAKEVGSEGKLGGQAQVRGVGGVWKDLTDNVNMMA
ANLTGQVRNIAEVTTAVARGDLSKKITVDVK-----GEILELKSTINTMVDQLNSFASEVTRVAREVGTEGKLGGQAQVEGVGGTWKDLTDNVNMMA
ANLTGQVRNIADVTTAVANGDLSKKITVDVK-----GEILELKNTINTMVDQLNSFASEVTRVAREVGSEGKLGGQARVEGVAGTWADLTDNVNLMA
ANLTGQVRNIADVTTAVANGDLSKKITVDVK-----GEILELKNTINTMVDQLNSFASEVTRVAREVGSEGKLGGQAQVRGVAGTWADLTDNVNLMA
ANLTGQVRNIAEVTTAVANGDLSKKITVDVK-----GEILELKSTINTMVDQLNSFASEVTRVAREVGTEGKLGGQAQVKGVGGVWKGLTDNVNMMA
ANLTGQVRNIAEVTTAVANGDLSKKITVAVE-----GEILELKSTINTMVDQLNSFASEVTRVAREVGIEGKLGGQAQVRGVGGTWKDLTDNVNMMA
ANLTGQVRNIADVTTAVANGDLSKKITVDVR-----GEILELKSTINTMVDQLNSFASEVTRVAREVGSEGKLGGQAQVKGVGGTWKDLTDNVNMMA
ANLTGQVRGIASVVTAVAQGDLKRKLTVDAK-----GEIAALADTVNEMIETLATFADQVTNVAREVGVEGKLGGQARVPGAAGLWRDLTDNVNQLA
ANLTTQVRAIADVATAVTKGDLARSISVEAS-----GEVASLKDTINEMIRNL
```