A Permutation Test for Stabilization of Polypeptide Helices by Sequence-Dependent Side Chain Interactions: Characterization of a Helix Initiation Site within the Myohemerythrin Sequence 76-87

Jonathan A. Zerkowski, Evan T. Powers, and D. S. Kemp*

Room 18-582, Department of Chemistry
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Only in unusual cases can one approximate the experimental helical propensity of an unaggregated peptide in water from its amino acid composition and sequence by using the conventional site and sequence independent initiation and propagation parameters.[1,2] Algorithms that rely solely on these paramaters fail if stabilizing side chain contacts of particular pairs of amino acids at $(i, i + 4)$ or $(i, i + 3)$ intraloop separations override sequence independent stability effects. Intraloop stabilizations of these types have been documented,[3] but no simple experimental test has previously measured their dominance within a particular peptide. If helicity is correctly modeled by sequence independent propagation constants (*s* values),[2] local sequence permutations must result in small helicity changes.[4] However, a large decrease in helicity is expected if a permutation disrupts stabilizing intraloop contacts. The peptide MHKDFLEKIGGL (*Themiste zostericola* myohemerythrin, residues 76-87) comprises the region of a larger peptide shown by Dyson et al. by CD and ^{1}H NMR analysis to be highly helical in trifluoroethanol−water mixtures.[5] We find that the helicity of this peptide is dramatically reduced by a permutation of the FLEKI sequence.

Trifluoroethanol (TFE) in dilute aqueous solution enhances helicity primarily by reducing the capacity of water to form hydrogen bonds with amides, increasing the helicity of peptides with a predisposition toward helix formation.[6] Consistent with
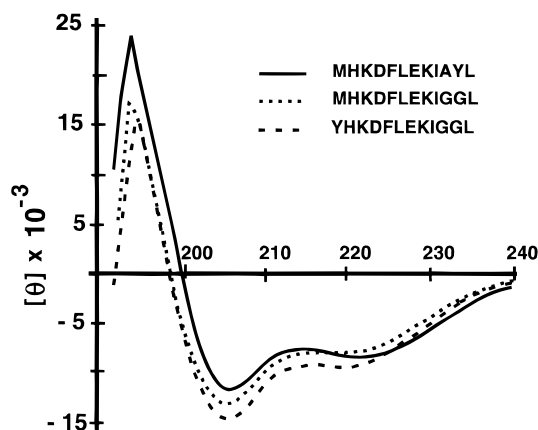
Figure 1. Circular dichroism (CD) spectra (mean residue molar ellipticity in deg cm^2 dmol^{-1} vs λ in nm) of sequence variants of the helical peptide MHKDFLEKIGGL at 5 °C, 10 mol % TFE−water, pH 10.3 (at which ellipticity is maximal). Replacement of the C-terminal sequence GGL by AYL, or M by Y at the N-terminus, results in small ellipticity changes. (Conversion of MHKDFLEKIAYL to Ac-MH-KDFLEKIAYL-NH$_2$ or to the sequence variant Ac-AHKDILEKYGGL-NH$_2$ results in no change in curve shape but enhances ellipticity by 1.7 and 1.3-fold, respectively.)

this effect, NMR spectroscopy of the myohemerythrin fragment in water alone reveals a disordered, protohelical structure.[5]

In 10 mol % TFE−water, the helical CD spectra of MH-KDFLEKIGGL and its variants YHKDFLEKIGGL and MH-KDFLEKIAYL are similar (Figure 1).[7] Switching Glu and Asp (MHKD̲F̲L̲E̲KIAYL → MHKE̲FL̲D̲KIAYL) likewise results in only a modest CD change (spectra not shown). The $(i, i + 4)$ spacing of an α-helix permits stabilizing hydrophobic contact between the side chains of Phe and Ile, similar to reported $(i, i + 4)$ Tyr-Val and Val-Tyr interactions.[3c,e] This stabilization is expected to be sustained by positional interchange, as well as by replacement of Phe by Tyr. For the benchmark MHKDFLE-KIAYL, a change to MHKDILEKF̲AYL (FLEKI → ILEKF) results in only a slight ellipticity change. However, the ellipticity is significantly reduced if Phe and Ile are replaced by Ala, widely viewed as a stronger helix stabilizer,[2] and the helical signature is abolished if the FLEKI sequence is permuted to EKFLI (Figure 2). This CD change implies a strong context dependence among the *s* values for amino acid residues at the FLEKI site.

Direct evidence for an $(i, i + 4)$ contact is provided by a crosspeak between the Phe aromatic and both Ile and Leu methyl resonances in a ROESY spectrum[8] of the helical analog YHDILEKFGGL (FLEKI → ILEKF), measured at 5 °C, pH 3.5 in D$_2$O-containing 10 mol % TFE-d_3. The assignment was confirmed by ROESY spectra of deuterium-labeled I and L peptides. ROESY spectra of the pentapeptide Ac-ILEKY-NH$_2$ under these conditions also show a crosspeak between Tyr aromatic and Ile + Leu methyl resonances, which is observed in D$_2$O alone or in 3:1 D$_2$O−DMSO-d_6, but at significantly reduced intensity. The CD spectrum of Ac-ILEKY-NH$_2$ (Figure 3) reflects nonrandom structure which is enhanced by TFE, abolished by the denaturant guanidine hydrochloride, and

* Author to whom correspondence should be addressed.
(1) Zimm, B. H.; Bragg, J. K. *J. Chem. Phys.* **1958**, *31*, 526−535. Lifson, S.; Roig, A. *J. Chem. Phys.* **1961**, *34*, 1963−1974. Poland, D.; Scheraga, H. A. *Statistical Mechanical Theory of Order-Disorder Transitions in Biological Macromolecules*; Academic Press: New York, 1970. Qian, H.; Schellman, J. A. *J. Phys. Chem.* **1992**, *96*, 3987−3994.
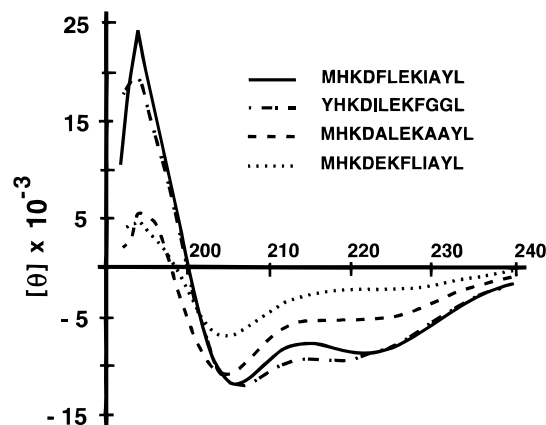(2) Wojcik, J.; Altmann, K.-H.; Scheraga, H. A. *Biopolymers* **1990**, *30*, 121−134. Chakrabartty, A.; Kortemme, T.; Baldwin, R. L. *Protein Sci.* **1994**, *3*, 843−852. Blaber, M.; Zhang, X. J.; Matthews, B. W. *Science* **1993**, *260*, 1637−1643. O'Neil, K. T.; DeGrado, W. F. *Science* **1990**, *250*, 646−651.
(3) (a) Vasquez, M.; Scheraga, H. A. *Biopolymers* **1988**, *27*, 41−58. (b) Schoemaker, K. R.; Fairman, R.; Schultz, D. A.; Robertson, A. D.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *Biopolymers* **1990**, *30*, 1−11. (c) Padmanabhan, S.; Baldwin, R. L. *J. Mol. Biol.* **1994**, *241*, 706−713. (d) Horovitz, A.; Matthews, J. M.; Fersht, A. R. *J. Mol. Biol.* **1992**, *227*, 560−568. (e) Nemethy, G.; Scheraga, H. A. *J. Phys. Chem.* **1962**, *66*, 1773−1789. (f) Loewenthal, R.; Sancho, J.; Fersht, A. R. *J. Mol. Biol.* **1992**, *224*, 759−770.
(4) Sets of literature *s* values[2] were applied to the helical part of the Lifson−Roig state sum[1d] as well as to the approximate state sum [helix state sum = $\sigma \sum_{j=1}^{(n-3)} (\sum_{i=0}^{(j-1)} (\Pi_{k=(j-i)}^{(n-1)} s_k))$] for which s_i is the *s* value at each site and *n* is the peptide length. For each set, the percentage change in helicity was calculated for the permutation MHKDFLEKIAYL → MH-KDEKFLIAYL, yielding an average change of 4.9% with a range of 0.2 to 9.7%; multiplication of all *s* values by 1.5 to approximate[6] the helicity increase in 10 mol % TFE gave a range of 0.1 to 7.8 %. These changes are clearly much smaller than the permutation-induced helicity changes of Figure 2. Applications of the helix state sum equation will be reported subsequently.
(5) Dyson, H. J.; Rance, M.; Houghten, R. A.; Wright, P. E.; Lerner, R. A. *J. Mol. Biol.* **1988**, *201*, 201−217. Dyson, H. J.; Merutka, G.; Waltho, J. P.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1992**, *226*, 790−817.
(6) (a) Cammers-Goodwin, A.; Allen, T. J.; Oslick, S. L.; McClure, K.; Lee, J. H.; Kemp, D. S. *J. Am. Chem. Soc.* **1996**, *118*, 3082−3090. (b) Lehrman, S. R.; Tuls, J. L.; Lund, M. *Biochemistry* **1990**, *29*, 5593−5596. (c) Segawa, S.-I.; Fukuno, T.; Fujiwara, K.; Noda, Y. *Biopolymers* **1991**, *31*, 499−509. (d) Conio, G.; Patrone, E.; Brighetti, S. *J. Biol. Chem.* **1970**, *245*, 3335−3340. (e) Storrs, R. W.; Truckses, D.; Wemmer, D. E. *Biopolymers* **1992**, *32*, 1695−1702.

(7) Peptides were synthesized, purified, and characterized as reported previously.[11a] CD spectra were obtained on an Aviv model 62DS in solutions buffered with 5 mM acetate/phosphate/borate in 5 mm or 1 mm quartz cells. Concentrations of tyrosine-containing peptides were determined using the UV absorption band at 275 nm; other peptides were quantitated by amino acid analysis. ^{1}H NMR spectra were recorded on a Varian VXR-501S spectrometer. ROESY-spectra[8] were acquired as reported previously,[9] using a 350 ms mixing time.
(8) Bothner-By, A. A.; Stephens, R. L.; Lee, J.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811−813.
(9) Groebke, K.; Renold, P.; Tsang, K.-W.; Allen, T. J.; McClure, K.; Kemp, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 2025−2029.

Figure 2. Circular dichroism (CD) spectra (mean residue molar ellipticity in deg cm^2 dmol^{-1} vs λ in nm) of sequence variants of the helical peptide MHKDFLEKIGGL at 5 °C, 10 mol % TFE−water, pH 10.3. Replacement of the FLEKI sequence by ALEKA or by EKFLI results in respective 1.6- and 3.9-fold decreases in negative ellipticity at 222 nm.
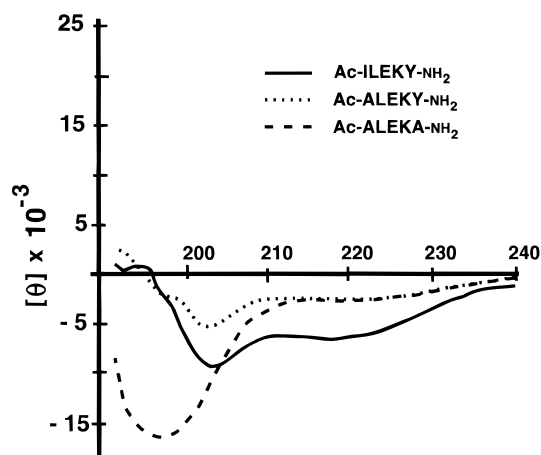
Figure 3. CD spectra (mean residue molar ellipticity in deg cm^2 dmol^{-1} vs λ in nm) of sequence variants of the pentapeptide sequence Ac-ILEKY-NH$_2$ at 5 °C, 10 mol % TFE−water, pH 10.3. Substitution of A for I to yield Ac-ALEKY-NH$_2$ reduces the negative ellipticity at 222 nm by 2.5-fold. The CD of Ac-ALEKA-NH$_2$ shows no evidence of nonrandom structure.

diminished for the sequence variants Ac-ALEKY-NH$_2$ and Ac-ALEKA-NH$_2$. The strong stabilization by TFE suggests the presence within the structure of intramolecular hydrogen bonding.[6] Several groups have recently reported similar short α- or 3$_{10}$-helical structures in the crystal and in nonaqueous solutions.[10]

We further characterized the relative helicities of these peptides by linking them to the reporting conformational template Ac-Hel$_1$. The ratio of s-trans to s-cis conformations of the acetyl group in Ac-Hel$_1$-peptide conjugates is proportional

to the peptide helicities.[11] In 4 mol % CF$_3$CD$_2$OD−D$_2$O at pH 10.5 and 25 °C, the four conjugates Ac-Hel$_1$-X-NH$_2$, where X = AFLEKIA, AILEKFA, AALEKAA, and AEKFLIA, show respective *t/c* ratios of 5.8, 4.4, 3.0, and 2.2, consistent with the order of 222 nm negative ellipticities observed for analogous peptides of Figures 2 and 3. These results show that a stabilizing interaction inferred for helical segments of proteins is mirrored in the properties of simple pentapeptides,[12] supporting a "bootstrap" analysis in which the complex energetic interactions characteristic of highly helical conformations of large peptides are modeled as a series of more tractable effects, observable in fragments of the larger segment.[13]

The dramatic decreases in helicity that result from modifications of the ($i, i + 4$) hydrophobic interactions of the FLEKI region suggest that the manifold of highly populated helical conformations of MHKDFLEKIGGL and its analogs is restricted to those that retain the Phe-Ile interaction or its equivalent. Two conceptually distinct models for the helical state sum are consistent with this inference. The conventional state sum for MHKDFLEKIGGL or its analogs is proportional to a constant, low-probability initiation parameter and contains terms proportional to the stability of every possible helical peptide conformer, including those that lack a helical FLEKI sequence. Multiplication of the terms corresponding to helical FLEKI-containing sequences by weighting factors would allow this state sum to mirror the stabilizing FLEKI interaction. Equivalently, exclusion of terms that lack this interaction would yield a much simpler state sum containing no weighting factors, but with a redefined, larger initiation parameter that reflects the high probability of helix nucleation from the FLEKI sequence. In TFE−water, helices formed from MHKDFLEKIGGL and its analogs are initiated by a single highly probable helical site consisting of the ($i, i + 4$) hydrophobic loop, first formed under conditions favorable to helix formation and last lost under denaturing conditions. It remains to be seen whether this case is unusual or whether other helical peptide sequences excised from proteins can be viewed as nucleated from a few highly probable initiation sites. The permutation test and the "bootstrap" analysis introduced in this paper should allow this question to be addressed experimentally.

(10) Vajayakumar, E. K. S.; Balaram, P. *Tetrahedron* **1983**, *39*, 2725−2731. Karle, I. L.; Balaram, P. *Biochemistry* **1990**, *29*, 6747−6756. Bobde, V.; Sasidhar, Y. U.; Durani, S. *Int. J. Pept. Protein Res.* **1994**, *43*, 209−218. Gallo, E. A.; Gellman, S. H. *J. Am. Chem. Soc.* **1994**, *116*, 11560−11561. Bobde, V.; Beri, S.; Durani, S. *Tetrahedron* **1993**, *49*, 5397−5406. Parthasarathy, R.; Chaturvedi, S.; Go, K. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 871−875.

(11) Kemp, D. S.; Allen, T. J.; Oslick, S. L. *J. Am. Chem. Soc.* **1995**, *117*, 6641−6657. Kemp, D. S.; Allen, T. J.; Oslick, S. L.; Boyd, J. G. *J. Am. Chem. Soc.* **1996**, *118*, 4240−4248. Kemp, D. S.; Oslick, S. L.; Allen, T. J. *J. Am. Chem. Soc.* **1996**, *118*, 4249−4255.

(12) The *t/c* value is proportional to the stability *S* of the frayed helical manifold initiated from the peptide−template junction: $t/c = A + BS$, where *A* and *B* are template-specific constants; $A = 0.8$.[11] The ratio of helical stabilities in 4 mol % TFE at 25 °C is $S_{AFLEKIA}/S_{AEKFLIA} = (5.8 − 0.8)/(2.2 − 0.8) = 3.6$.

(13) A referee has pointed out that the α-helix-stabilizing ($i, i + 4$) FLEKI interaction may reduce the otherwise potentially significant 3$_{10}$ character in these short helices[14] and that the CD spectrum of the alanine-rich MHKDALEKAAYL (Figure 2) is not inconsistent with presence of weak 3$_{10}$ helix.[15]

(14) Fiori, W. R.; Millhauser, G. L. *Biopolymers* **1995**, *37*, 243−250.

(15) Toniolo, C.; Polese, A.; Formaggio, F.; Crisma, M.; Kamphuis, J. *J. Am. Chem.* Soc. **1996**, *118*, 2744−2745.