Origin of the p K a Perturbation of N-Terminal Cysteine in α- and 3 10 -Helices: A Computational DFT Study

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · FEBRUARY 2006

Impact Factor: 3.3 · DOI: 10.1021/jp0549780 · Source: PubMed

3 AUTHORS:

Goedele Roos
Vrije Universiteit Brussel
41 PUBLICATIONS **587** CITATIONS
SEE PROFILE

Stefan Loverix
Mabmole
27 PUBLICATIONS **814** CITATIONS
SEE PROFILE

Paul Geerlings
Vrije Universiteit Brussel
460 PUBLICATIONS **11,544** CITATIONS
SEE PROFILE

Origin of the pK_a Perturbation of N-Terminal Cysteine in α- and 3_{10}-Helices: A Computational DFT Study

Goedele Roos,[†] Stefan Loverix,[§,†] and Paul Geerlings*[,†]

Algemene Chemie (ALGC) and Departement Ultrastructuur, Vlaams Interuniversitair Instituut voor Biotechnologie (VIB), Vrije Universiteit Brussel (VUB), Pleinlaan 2, B-1050, Brussels, Belgium

Received: September 2, 2005; In Final Form: October 12, 2005

It is well documented that helices in proteins can decrease the pK_a of residues located at the N-terminus, but the real nature of this perturbation remains unclear. In the present work, the origin of the effect of 3_{10}- and α-polyalanine helices on the pK_a of an N-terminal cysteine residue is examined in gas phase as well as in aqueous solution by means of density functional theory. In a systematic study of the helix dipole, the proton affinity (PA), and the pK_a of the N-terminal cysteine, in relation to both the helix length and the strength of the hydrogen bonds between the helix backbone amides and the Sγ of the N-terminal cysteine, a direct relation between the terminal hydrogen bonds and the pK_a perturbation is revealed.

Introduction

Elements of secondary structure such as 4_{13}- (or α-) and 3_{10}-helices are ubiquitous and important structural features in proteins.[1−3] α-Helices are the most common type of secondary structures, while 3_{10}-helices are the fourth common type. Most α-helices in proteins contain 10−15 residues,[4] while 3_{10}-helices usually form short sequences of 4−6 amino acids.[4] The symbols 3_{10} and 4_{13} imply that the intramolecular hydrogen bonds between the backbone carbonyl oxygens of residue *i* and the amide protons of residue $i + 3$ or $i + 4$ form a ring containing 3 or 4 sequential carbonyl oxygens, consisting of 10 or 13 atoms, respectively. Consequently, for the same number of amino acids, 3_{10}-helices have the most hydrogen bonds. Nevertheless, the α-helix is dominant in protein structures.[5−7]

The helix macrodipole is the vector sum of the microdipole moments of the individual peptide units and is oriented along the helix axis. A single peptide unit has a considerable dipole moment because of the partial double-bond character of the N−C bond.[7,8] A commonly accepted value for this dipole moment is 3.5 D.[9] The direction of the dipole is parallel to the C═O and N−H bonds. (Figure 1).

In an α-helix, the peptide units are aligned in such a manner that ∼97% of the peptide dipole moments point in the direction of the helix axis.[10] The 3_{10}-conformation is disfavored because the H-bond geometry is not as optimal, leading to a less optimal orientation of the microdipoles with respect to the 3_{10}-helix axis.[6] The effect of the α-helix dipole has been suggested to be equivalent to that of a −0.5 unit charge at the C-terminus plus a +0.5 unit charge at the N-terminus of the helix.[9,11] The dipolar nature of α-helices has been invoked to explain the structure of ligand-binding sites,[9] the relative disposition of α-helices in proteins,[12,13] and the clustering of positive and negative charges toward the C- and N-termini of the helices.[14] The interaction between α-helix dipoles and charged residues in small peptides[15,16] as well as in proteins[17−19] has been acknowledged to contribute to stability.

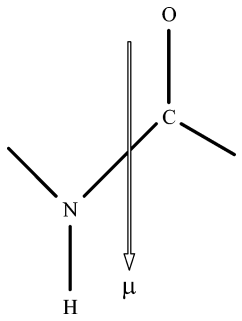

Figure 1. Dipole of one peptide unit.

It is both experimentally and computationally well documented that helices can influence the pK_a of residues located at either the N- or C-termini. Experimentally, a pK_a increase of 0.6, 1.6, and 2.2 units was found for the C-terminal histidine residue in triosephosphate isomerase,[21] barnase,[20] and human hemoglobin, respectively.[17] Similar studies on an N-terminal cysteine residue found a pK_a decrease of 1.8 and 2.0 units in rhodanese[22] and human thioredoxin,[23] respectively. The pK_a of N-terminal aspartate in an experimentally designed dodecapeptide is suppressed by 0.6 units.[25] Earlier quantum chemical studies on papain have shown that a helix near the active site facilitates the proton transfer from the N-terminal serine to a catalytic histidine.[26] A previous density functional theory (DFT) study (including electron correlation) of our group on the N-terminal Cys10 in arsenate reductase[24] suggested a pK_a decrease of 2.3 units.

The macrodipole concept has been reexamined several times in the literature. Because of the often-used term "helix macrodipole", which might have given the impression that both ends of the helix contribute significantly to the overall effect, the short-range nature of the helix effect has not been appreciated for a long time.[27] Electrostatic free energy calculations suggest that the first turn of the helix (e.g., by providing hydrogen bonds) accounts for about 80% of the overall charge-stabilization effect,[27] while a mutagenesis study gives further indication of the importance of the terminal hydrogen bonds in pK_a perturba-

* Corresponding author. Phone: +32 2 629 33 14. Fax: +32 2 629 33 17. E-mail: pgeerlin@vub.ac.be.

[†] Algemene Chemie (ALGC).

[§] Departement Ultrastructuur, Vlaams Interuniversitair Instituut voor Biotechnologie (VIB).

10.1021/jp0549780 CCC: $33.50
Published on Web 12/08/2005

TABLE 1: Ψ and Φ Torsional Angles of Fully Optimized α-Helices (Intermediate 2) and 3_{10}-Helices (Intermediate 1) of (A) Four and (B) Six Amino Acids

	(A)	
Intermediate 1	Φ_1: −120	Ψ_1: 16
	Φ_2: −68	Ψ_2: −22
	Φ_3: −70	Ψ_3: −8
	Φ_4: −104	Ψ_4: −11
Intermediate 2	Φ_1: −121	Ψ_1: 16
	Φ_2: −68	Ψ_2: −22
	Φ_3: −68	Ψ_3: −8
	Φ_4: −104	Ψ_4: −11
S_Intermediate 1	Φ_1: −74	Ψ_1: −79
	Φ_2: −64	Ψ_2: −23
	Φ_3: −67	Ψ_3: −11
	Φ_4: −98	Ψ_4: 7
S_Intermediate 2	Φ_1: −78	Ψ_1: −75
	Φ_2: −68	Ψ_2: −15
	Φ_3: −79	Ψ_3: −14
	Φ_4: −115	Ψ_4: −27
	(B)	
Intermediate 1	Φ_1: −114	Ψ_1: 9
	Φ_2: −65	Ψ_2: −26
	Φ_3: −62	Ψ_3: −18
	Φ_4: −63	Ψ_4: −20
	Φ_5: −71	Ψ_5: −6
	Φ_6: −103	Ψ_6: −11
Intermediate 2	Φ_1: −113	Ψ_1: 9
	Φ_2: −65	Ψ_2: −26
	Φ_3: −62	Ψ_3: −18
	Φ_4: −63	Ψ_4: −19
	Φ_5: −70	Ψ_5: −8
	Φ_6: −103	Ψ_6: −11
S_Intermediate 1	Φ_1: −75	Ψ_1: −83
	Φ_2: −62	Ψ_2: −25
	Φ_3: −59	Ψ_3: −22
	Φ_4: −63	Ψ_4: −21
	Φ_5: −70	Ψ_5: −8
	Φ_6: −99	Ψ_6: 7
S_Intermediate 2	Φ_1: −77	Ψ_1: 77
	Φ_2: −60	Ψ_2: −31
	Φ_3: −63	Ψ_3: −33
	Φ_4: −77	Ψ_4: −37
	Φ_5: −62	Ψ_5: −24
	Φ_6: −77	Ψ_6: −11

tion.[11] Similarly, a Hartree−Fock study on papain attributes more than half of the helical effect to hydrogen bonds with the backbone rather than to the macrodipole.[28]

Most of the experimental research on the helical influence on pK_a dates from the 1980s and early 1990s. At that time, high-level quantum chemical computational research on large systems such as helices was not possible. The ever-increasing computational power has made quantum chemical studies on biosystems achievable today. Such studies have the advantage that they can shed light on problems that are experimentally difficult to access.[29]

In the present work, the effect of 3_{10}- and α-polyalanine helices on the properties of an N-terminal cysteine residue is examined in gas phase and in aqueous solution. Our major objective is to investigate the individual roles of the hydrogen-bond pattern with the N-terminal cysteine residue and of the helix dipole in pK_a perturbation. Hence, DFT calculations of the helix dipole, the proton affinity (PA), and the pK_a of the N-terminal cysteine were carried out as functions of both the helix length and the strength of the hydrogen bonds between the helix backbone amides and the Sγ of the N-terminal cysteine.

Computational Details

$(Ala)_n$ (with $n = 2, 3, 4, 6, 8$) and Cys1-$(Ala)_n$ (with $n = 1, 2, 3, 5, 7$) polypeptide chains with an α- and 3_{10}-helical conformations were constructed. The torsional angle Φ defines the rotation of the plane containing C^{α}_i, C'_i and O_i (and N_{i+1}) around the N_i-C^{α}_i bond, while the angle Ψ defines the rotation of the plane containing C'_i and O_i and N_{i+1} around the C^{α}_i−C_i bond. Right-handed α-helices have typical torsion angles of $\Psi = -57$ degrees and $\Phi = -47$ degrees, whereas 3_{10}-helices have torsion angles of $\Psi = -49$ degrees and $\Phi = -26$ degrees.

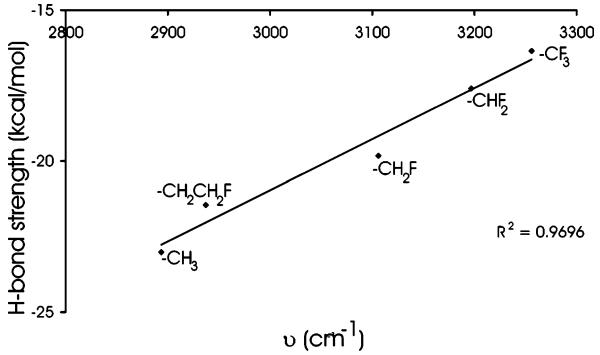

Figure 2. Stretching frequencies of the bond between the hydrogen donor (nitrogen atom) and the hydrogen, and the hydrogen bond strength (calculated at the B3LYP/6-31+G** level with the CP method[28] to correct for BSSEs) of five CH_3COCH_3NH−SX (with X = $-CH_3$, $-CH_2F$, $-CHF_2$, $-CHF_3$, and $-CH_2CH_2F$) hydrogen bonds.

Full relaxation of the native geometries of both true α- and 3_{10}-helical forms converges to very similar intermediate structures, as judged by their dihedral angles (Table 1). This is especially true for $n = 4$ and in accordance with ref 5. When the number of amino acids is increased or in the presence of the N-terminal cysteine, the geometries of both optimized structures diverge more, but the optimized dihedral Ψ and Φ angles deviate significantly from their starting values (Table 1). Therefore, during the optimization, the torsional Ψ and Φ angles were kept fixed to ensure the desired helix type was preserved.

For convenience, the helices are named accordingly: type of helix + number of amino acids; when the N-terminal residue is a cysteine, an S is put in front of the name: for example, α-4 is an α-helix composed of 4 amino acids; S-α-4 is an α-helix composed of 4 amino acids of which the first is a cysteine. Helices obtained after full optimization starting from the 3_{10}- and α-helical forms are named "Intermediate 1" and "Intermediate 2", respectively.

Since dipoles of charged residues are origin dependent, the N-terminal nitrogen atom was chosen as the origin to ensure a meaningful comparison between the helix dipoles in the presence of a charged N-terminal cysteine. The current computational power allows for the calculation of the helix dipole for the system as a whole, rather than as the sum of monomeric contributions, taking into account interactions among residues governed among others by the polarizing effect of one monomer on another.

To translate changes in −NH proton-donor stretching frequencies to changes in H-bond strength, we calculated the stretching frequencies of a series of five CH_3−CO−NH(- - - SX)−CH_3 (with X = $-CH_3$, $-CH_2F$, $-CHF_2$, $-CHF_3$, and $-CH_2CH_2F$) hydrogen bonds (Figure 2) and plotted these values against the calculated hydrogen-bond strengths. The basis set superposition error (BSSE), was taken into account by the counterpoise (CP) correction proposed by Boys and Bernardi.[30] The CH_3−CO−NH(- - - SX)−CH_3 structures were fully optimized (Figure 3).

To translate changes in natural population analysis (NPA) charge to changes in the acid dissociation constant (pK_a), we

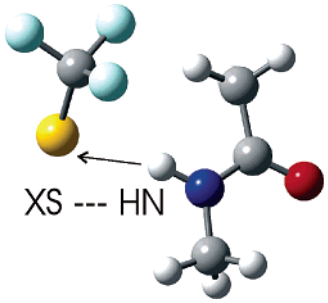

Figure 3. $CH_3-CO-NH(- - - SX)-CH_3$ (with X = $-CH_3$, $-CH_2F$, $-CHF_2$, $-CHF_3$, and $-CH_2CH_2F$) hydrogen bonds. Color Code: yellow = sulfur; blue = nitrogen; white = hydrogen; red = oxygen; gray = carbon; and cyan = fluorine.

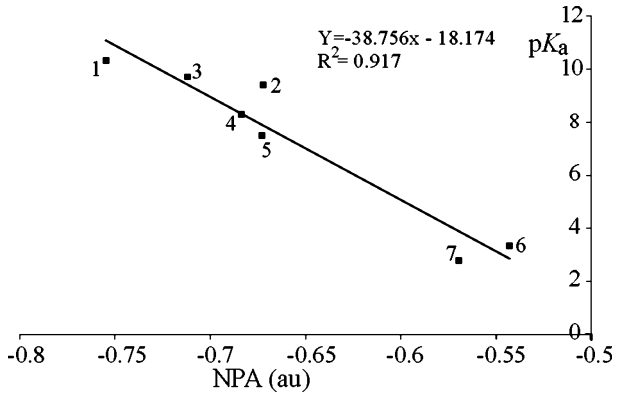

Figure 4. NPA-charge (of the sulfur atoms of the thiolates)−pK_a calibration curve obtained for a series of substituted thiolates (methanethiol (1), benzenemethanethiol (2), mercaptoethanol (3), cysteine (4), trifluoroethanethiol (5), thioacetic acid (6), and trifluoromethanethiol (7)) in gas phase.

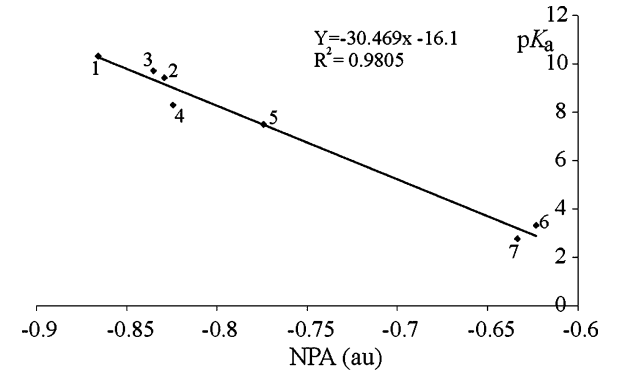

Figure 5. NPA-charge (of the sulfur atoms of the thiolates)−pK_a calibration curve obtained for a series of substituted thiolates (methanethiol (1), benzenemethanethiol (2), mercaptoethanol (3), cysteine (4), trifluoroethanethiol (5), thioacetic acid (6), and trifluoromethanethiol (7)) in aqueous solution.

calculated the NPA charges on the sulfur atom of a series of seven thiolates (methanethiol, benzenemethanethiol, mercaptoethanol, cysteine, trifluoroethanethiol, thioacetic acid, and trifluoromethanethiol) in gas phase as well as in aqueous solution and plotted these values against experimental pK_a values (Figures 4 and 5). The resulting linear relationship was used to extrapolate the pK_a of the N-terminal cysteine in the helices from the calculated NPA charge of the N-terminal SγCys atom. The structures of the thiolates were fully optimized.

To account for solvent effects, the polarizable continuum model (PCM)[31] solvent model was used, implemented in the Gaussian 03 package.[32] All optimizations were performed at the B3LYP level using the 6-31G* basis set, and subsequent calculations were performed at the B3LYP level with the 6-31+G** basis set. The Gaussian 03 package[32] was used throughout.

Results and Discussion

The N-terminal cysteine accepts two hydrogen bonds from nearby helix backbone amides. The lengths of these hydrogen bonds are smaller in the α- than they are in the 3_{10}-helices for an equal number of residues (Table 2). When the number of amino acids of the helix increases, the hydrogen-bond length slightly decreases.

To give a quantitative assessment of the intramolecular hydrogen bonds, the stretching frequency of the NH bond involved in the hydrogen bonding was calculated. Upon the formation of a hydrogen bond, a red shift in this NH stretching frequency occurs. To translate the magnitude of this red shift to the hydrogen bond strength, a series of $CH_3-CO-NH(- - - SX)-CH_3$ hydrogen-bonded model systems was used. In this series, plotting the calculated H-bond strength versus the calculated NH bond-stretching frequency yields a linear correlation (Figure 2). This curve was used to estimate the hydrogen bond strength in the helices from the calculated stretching frequencies (Table 2).

The results in Table 2 show that for both the α- and 3_{10}-helices, the hydrogen bond strength increases with an increasing number of amino acids. These results also indicate that the hydrogen bonds between the helix backbone and SγCys are stronger in the α-helices compared to those in the 3_{10}-helices.

The macrodipole increases with the number of amino acids (and consequently with the number of microdipoles) (Table 3). The less optimal orientation of the hydrogen bonds and microdipoles in a 3_{10}-helix can explain the lower value of the macrodipole of a 3_{10}-helix compared to that in an α-helix of the same length.[6] The dipoles of the fully optimized 3_{10}- and α-helices (Intermediate 1 and Intermediate 2) of the same length are only very slightly different (Table 3), illustrating the convergence of structures during a full optimization (see also Table 1).

When the N-terminal cysteine is present, the value of the macrodipole increases because of the electrostatic interaction between the charged sulfur atom and the helix backbone (Table 3). This interaction seems to cause a better conservation of the initial helical structure during a full optimization.

Surprisingly, in the presence of a N-terminal cysteine, the α-helix macrodipole is smaller than that of a 3_{10}-helix. The hydrogen bonds formed between the α-helical backbone and SγCys are stronger than these formed between the 3_{10}-helical backbone and SγCys (Table 2). It would be interesting to know to what extent the dipole effect and the terminal hydrogen bonds contribute to the pK_a perturbation of the cysteine.

From Table 4, it can be seen that the PA of the N-terminal cysteine decreases when the number of amino acids increases. In parallel, the macrodipole increases, as does the strength of the backbone amide−SγCys hydrogen bonds.

For helices of the same length, the PA of the N-terminal cysteine is lower in α-helices than it is in 3_{10}-helices. This observation can be explained by the stronger backbone amide−SγCys hydrogen bonds found in the α-helices (Table 2). Although the dipole moment—and its interaction with the charged cysteine residue—is larger in the 3_{10}-helix, the net effect on the PA is larger in the α-helix. This indicates that it is mainly the hydrogen bonds with SγCys, which influence the PA, and, by extension, the pK_a, since a relation exists between PA and pK_a.[33]

TABLE 2: Length (*l*) and Donor−Hydrogen-Acceptor Angle *a* (°) of the Hydrogen Bonds Formed between SγCys and the Helix Backbone, and Stretching Frequencies ν (cm^{-1}) and Hydrogen-Bond Strengths (kcal/mol) of the Proton-Donor NH-Bond as a Function of the Number of Residues

(A) Pure 3_{10}- and α-Helices						(B) Intermediate Structures[c]					
	donor−acceptor	*l* (Å)[a]	*a* (°)	ν (cm^{-1})	H-bond strength (kcal/mol)[b]		donor−acceptor	*l* (Å)[a]	*a* (°)	ν (cm^{-1})	H-bond strength (kcal/mol)[b]
		2 AA						4 AA			
S-3_{10}	$NH_{N-terminus}$−S	3.010	119	3277	−16.3	S_Intermediate 1	$NH_{N-terminus}$−S	3.080	116	3388	−14.5
	NH_1−S	3.410	125	3405	−14.2		NH_1−S	3.099	149	2892	−22.8
S-α	$NH_{N-terminus}$−S	3.040	119	3329	−15.5	S_Intermediate 2	$NH_{N-terminus}$−S	3.090	116	3383	−14.5
	NH_1−S	3.200	142	3138	−18.7		NH_1−S	3.090	149	2856	−23.5
		3 AA						6 AA			
S-3_{10}	$NH_{N-terminus}$−S	3.010	120	3272	−16.4	S_Intermediate 1	$NH_{N-terminus}$−S	3.100	116	3388	−14.5
	NH_1−S	3.370	126	3378	−14.6		NH_1−S	3.070	149	2838	−23.8
S-α	$NH_{N-terminus}$−S	3.043	119	3336	−15.3	S_Intermediate 2	$NH_{N-terminus}$−S	3.090	116	3381	−14.6
	NH_1−S	3.190	144	3068	−19.9		NH_1−S	3.060	149	2798	−24.4
		4 AA									
S-3_{10}	$NH_{N-terminus}$−S	3.010	119	3272	−16.4						
	NH_1−S	3.340	125	3320	−15.6						
S-α	$NH_{N-terminus}$−S	3.040	119	3322	−15.6						
	NH_1−S	3.160	142	3001	−21.0						
		6 AA									
S-3_{10}	$NH_{N-terminus}$−S	3.010	118	3282	−16.3						
	NH_1−S	3.310	128	3303	−15.9						
S-α	$NH_{N-terminus}$−S	3.034	119	3308	−15.8						
	NH_1−S	3.130	145	2929	−22.2						
		8 AA									
S-3_{10}	$NH_{N-terminus}$−S	3.012	199	3270	−16.5						
	NH_1−S	3.290	128	3293	−16.1						
S-α	$NH_{N-terminus}$−S	3.032	119	3305	−15.9						
	NH_1−S	3.122	145	2891	−22.9						

[a] Distance measured in Å from H-donor to H-acceptor. [b] Hydrogen bond strengths obtained via the linear relationship of Figure 2 from the stretching frequencies of the bond between the hydrogen donor and the hydrogen.

TABLE 3: Macrodipole Moments of the Different Helical Types with Different Lengths

dipole (D) B3LYP/6-31G*	1 AA	2 AA	3 AA	4 AA	6 AA	8 AA
alanine	3.6407					
cysteine	4.3963					
3_{10}		3.60	7.74	10.47	18.70	
α		4.78	8.20	11.88	19.26	
S_3_{10}		10.63	12.28	16.01	24.50	34.33
S_α		9.67	11.29	13.18	23.32	32.39
Intermediate 1				8.23	16.33	
Intermediate 2				8.22	16.38	
S_Intermediate 1				10.72	20.32	
S_Intermediate 2				11.87	21.16	

Among others, the NPA charge has been shown to be an effective descriptor for the pK_a.[24,33] In a series of thiolates, a linear relationship is obtained between the NPA-charge of the sulfur atom and the experimental pK_a value (Figure 4). The more negative the NPA charge on the sulfur atom, the higher the tendency to bind a proton, and as a result, the more basic (i.e., higher pK_a) the compound is. This linear relationship can be used as a calibration curve to quantify the pK_a perturbing effect.

When calibrating the calculated NPA charges of the SγN-terminal Cys in helices, an increase (less negative) in the NPA charge and, by consequence, a pK_a decrease with the number of amino acids is found. For helices of the same length, a lower pK_a value of the N-terminal cysteine residue is found in the α-helix (Table 4) compared to that found in the 3_{10}-helix. This is in accordance with the trends indicated for the PA and in agreement with the higher hydrogen-bond strengths found for hydrogen bonds with the N-terminal SγCys atom in the α-helices compared to those found in the 3_{10}-helices.

TABLE 4: PA, NPA Charge, and pK_a of N-Terminal SγCys in (A) 3_{10}-Helices, (B) α-Helices, and (C) Intermediate Structures

(A)

	PA Sγ (au)	NPA Sγ (au)	pK_a	additional pK_a decrease per amino acid
cysteine	−0.556	−0.684	8.3 (exp.)	
S_3_{10}_2	−0.528	−0.642	6.7	−1.6
S_3_{10}_3	−0.519	−0.633	6.3	−0.4
S_3_{10}_4	−0.514	−0.626	6.1	−0.3
S_3_{10}_6	−0.503	−0.621	5.9	−0.2
S_3_{10}_8		−0.616	5.7	−0.2

(B)

	PA Sγ (au)	NPA Sγ (au)	pK_a	additional pK_a decrease per amino acid
S_α_2	−0.519	−0.623	6.0	−2.3
S_α_3	−0.510	−0.610	5.5	−0.5
S_ α_4	−0.502	−0.600	5.1	−0.4
S_α_6	−0.478	−0.586	4.5	−0.6
S_α_8		−0.578	4.2	−0.3

(C)

	PA Sγ (au)	NPA Sγ (au)	pK_a
S_Intermediate 1_4	−0.508	−0.601	5.1
S_Intermediate 1_6	−0.496	−0.593	4.8
S_Intermediate 2_4	−0.501	−0.594	4.9
S_Intermediate 2_6	−0.497	−0.584	4.5

Compared to isolated cysteine, the addition of one extra amino acid causes the formation of two hydrogen bonds between the backbone amides and SγCys (Table 2). These two hydrogen

TABLE 5: Helix Macrodipole, NPA Charge, and pK_a of N-terminal SγCys Obtained in Aqueous Solution in (A) 3_{10}-Helices and (B) α-Helices

(A)

	dipole (D)	NPA Sγ (au)	pK_a	additional pK_a decrease per amino acid
cysteine	6.45	−0.824	8.3 (exp.)	
S_3_{10}_2	14.33	−0.791	8.0	−0.3
S_3_{10}_3	14.38	−0.786	7.9	−0.1
S_3_{10}_4	17.97	−0.782	7.8	−0.1

(B)

	dipole (D)	NPA Sγ (au)	pK_a	additional pK_a decrease per amino acid
S_α_2	11.64	−0.763	7.2	−1.1
S_α_3	12.77	−0.758	7.0	−0.2
S_ α_4	15.32	−0.753	6.9	−0.1
S_α_6	26.65	−0.754	6.9	+0.0

bonds are responsible for a substantial decrease in pK_a (1.9 units in 3_{10}-helices and 2.3 units in α-helices) (Table 4). The addition of a third, fourth, and so on residue in the conformation of a 3_{10}- or α-helix results in the strengthening of these hydrogen bonds (Table 2). The more residues the helix counts, the larger the macrodipole, and apparently the stronger the hydrogen bonds to SγCys are. The latter may be expected from the electrostatic nature of hydrogen bonds. Per extra residue, a further pK_a decrease is found, however, to a lesser extent (Table 4). This decrease diminishes for every additional residue, leading to a plateau value. As a result, additional residues in a helix strengthen the terminal hydrogen bonds, but cause a subordinate effect on pK_a. Interestingly, the macrodipole increases linearly with the number of residues (Table 3), in contrast to the pK_a effect, suggesting a subordinate effect of the helix macrodipole.

The pK_a and PA values of the N-terminal cysteines present in the intermediate helical structures obtained after full optimization (Intermediate 1 and Intermediate 2) are lower than those in 3_{10}- or α-helices with the same number of amino acids. The hydrogen bonds formed between the backbone amides and the SγCys found in Intermediate 1 and Intermediate 2 are stronger than those found in the 3_{10}- and α-helices of the same length (Table 2), as could be expected. On the other hand, the macrodipoles of the intermediates are smaller than those of 3_{10}- or α-helices (Table 3). Here again, the terminal hydrogen bonds have a major influence on the pK_a perturbation.

Because helices in biological systems (e.g., proteins) are exposed mostly to solvent, we will also use a solvent model to test the validity of the observations in gas phase. Therefore, the S-α- and S-3_{10}-helices were optimized with fixed Ψ and Φ (vide supra) in water, using a PCM model.[30]

All results obtained in gas phase remain valid for the solvated helices. For helices of the same length, the macrodipoles of the solvated 3_{10}-helices are larger than these of the solvated α-helices, while the NPA charge of the N-terminal SγCys atom present in the α-helices is higher (less negative) than that of the SγCys in the 3_{10}-helices (Table 5). As was done in gas phase, a linear relationship between the NPA charges calculated in an aqueous solution of the thiolate−sulfur atoms of the series of substituted thiolates and their experimental pK_a value is obtained (Figure 4). When calibrating the NPA charge of the SγCys of the solvated helices, a decrease in the pK_a with the helical length is found. As was the case in gas phase, this decrease is accompanied by an increase in the backbone amide−SγCys hydrogen-bond strength as measured from frequency calculations (Table 6).

The extra decrease in the pK_a for every additional residue diminishes, which means that once the hydrogen bonds are formed (when residue 2 is present), the next residues, which strengthen the hydrogen bonds, cause a subordinate effect on pK_a perturbation. The very small pK_a increase of 0.03 units found in the presence of a six-residue α-helix is likely to be due to reaching the plateau pK_a value, rather than representing a relevant effect.

For helices with the same length, the dipoles in water are larger than those in gas phase.

In gas phase, the pK_a value of 4.2 obtained for a cysteine residue present at the N-terminal of an α-helix with eight residues is too low in comparison with literature values.[22−24] On the other hand, in aqueous solution, a decrease in the N-terminal cysteine pK_a of 1.4 units due to a six-residue α-helix is found. This is in the range of pK_a values of N-terminal cysteines obtained by experimental studies.[22,23] This more realistic pK_a decrease calculated in aqueous solution is consistent with the diminished electrostatic effect of the hydrogen bonds in solvent in comparison to that in gas phase.

Although 3_{10}-helices are very important structural features in protein functioning (e.g., in cycline-dependent kinase[34] and in HIV antibodies[35]), to the best of our knowledge, no role in pK_a perturbation has been described yet for 3^{10}-helices. The present theoretical study shows that 3^{10}-helices can indeed lower the pK_a, albeit to a lesser degree than can α-helices.

The calculations in gas phase and in solvent point out that the pK_a lowering effect on the N-terminal cysteine of an α- or 3_{10}-helix largely finds its origin in the terminal hydrogen bonds formed with this N-terminal residue. These results strengthen the view of the short-range nature of the helical effect on pK_a perturbation and the significance of the terminal hydrogen bonds herein, as was proposed previously from electrostatic free energy calculations,[27] mutagenesis studies,[11] and quantum chemical studies.[28] There is no reason the results of the present study on a N-terminal cysteine residue would not be valid for other N- or C-terminal residues, provided that hydrogen bonds with donor or acceptor atoms of the helix can be formed. However, the extension of the present results to systems with N- or C-terminal residues different from cysteine needs further confirmation. In the absence of such hydrogen bonds, the macrodipole effect

TABLE 6: Length (l)[a] and Donor−Hydrogen-Acceptor Angle a (°) of the Hydrogen Bonds Formed between SγCys and the Helix Backbone in Aqueous Solution and Stretching Frequencies ν (cm^{-1}) of the Proton-Donor NH-Bond

		2 AA			3AA			4AA			6AA		
	donor−acceptor	l (Å)	a (°)	ν (cm^{-1})	l (Å)	a (°)	ν (cm^{-1})	l (Å)	a (°)	ν (cm^{-1})	l (Å)	a (°)	ν (cm^{-1})
S-3_{10}	$NH_{N-terminus}$−S	3.190	110	3310	3.178	111	3353	3.170	108	3316			
	NH_1−S	3.696	104	3504	3.681	105	3464	3.690	110	3503			
S-α	$NH_{N-terminus}$−S	3.133	112	3318	3.135	112	3314	3.141	112	3313	3.170	110	3315
	NH_1−S	3.357	131	3427	3.342	131	3404	3.328	131	3406	3.327	129	3372

[a] Distance measured in Å from H-donor to H-acceptor.

alone is responsible for charge stabilization at these residues, as has been concluded before for the Cys−His ion pair in papain.[36]

Conclusion

We conclude from this study that the dominant pK_a perturbing effect of helices on an N-terminal cysteine largely finds its origin in the terminal hydrogen bonds, which are strengthened as the helix length increases with every extra amino acid added. Each additional residue has a subordinate effect on the pK_a, which may lead to a final plateau value of the pK_a. Similar trends are found in gas phase and in solvent, but those in the latter are found to be necessary to obtain reliable values for the pK_a decrease.

Acknowledgment. P.G. wants to thank the Fund for Scientific Research Flanders (FWO) and the VUB for continuous support of his research group. S.L. and G.R. thank the Fund for Scientific Research Flanders (FWO) for a postdoctoral and predoctoral fellowship (Aspirant), respectively.

References and Notes

(1) (a) Richardson, J. S. *Adv. Protein Chem.* **1981**, *34*, 168. (b) Creighton, T. E. *Proteins: Structures and Molecular Properties*, 2nd ed.; Freeman: New York, 1993.

(2) Hunter, T.; Pines, J. *Cell* **1994**, *79*, 573.

(3) Hashimoto, Y.; Kohri, K.; Kaneko, Y.; Morisaki, H.; Kato, T.; Ikeda, K.; Nakanishi, M. *J. Biol. Chem.* **1998**, *273*, 16544.

(4) (a) Wu, Y.-D.; Zhao, Y.-L. *J. Am. Chem. Soc.* **2001**, *123*, 5313. (b) Pal, L.; Basu, G.; Chakrabarti, P. *Proteins: Struct., Funct., Genet.* **2002**, *48*, 571.

(5) Topol, I. A.; Burt, S. K.; Deretey, E.; Tang, T.-H.; Perczel, A.; Rashin, A.; Csizmadia, I. G. *J. Am. Chem. Soc.* **2001**, *123*, 6054.

(6) Tran, T. T.; Zeng, J.; Treutlein, H.; Burgess, A. W. *J. Am. Chem. Soc.* **2002**, *124*, 5222.

(7) Schulz, G. E.; Schirmer, R. H. *Principles of Protein Structure*; Springer-Verlag: New York, 1979.

(8) Pauling, L. *The Nature of the Chemical Bond*, Cornell University Press: Ithaca, New York, 1960.

(9) Hol, W. G. J.; Van Duijnen, P. T.; Berendsen, H. J. C. *Nature* **1978**, *294*, 443.

(10) Wada, A. *Adv. Biophys.* **1976**, *9*, 1.

(11) Sancho, J.; Serrano, L.; Fersht, A. R. *Biochemistry* **1992**, *31*, 22.

(12) Sheridan, R. P.; Levy, R. M.; Salemme, F. R. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4545.

(13) Presnell, S. R.; Cohen, F. E. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 6592.

(14) Richardson, J. S.; Richardson, D. C. *Science* **1988**, *240*, 1648.

(15) Shoemaker, K. R.; Kim, P. S.; Brems, D. N.; Marqusee, S.; York, E. J.; Chiken, I. M.; Steward, J. M.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 2349.

(16) Fairman, R.; Shoemaker, K. R.; York, E. J.; Steward, J. M.; Baldwin, R. L. *Proteins: Struct., Funct., Genet.* **1989**, *5*, 1.

(17) Perutz, M. F.; Gronenborn, A. M.; Clore, G. M.; Fogg, J. H.; Shih, D. T.-b. *J. Mol. Biol.* **1985**, *183*, 491.

(18) Sali, D.; Bycroft, M.; Fersht, A. R. *Nature* **1988**, *335*, 740.

(19) Nicholson, H.; Becktel, W. J.; Matthews, B. W. *Nature* **1988**, *336*, 651.

(20) Sali, D.; Fersht, A. R.; Bycroft M. *Nature* **1988**, *335*, 6192.

(21) Lodi, P. J.; Knowles, J. R. *Biochemistry* **1993**, *32*, 4338.

(22) Schlesinger, P.; Westley, J. *J. Biol. Chem.* **1973**, *240*, 780.

(23) Forman-Kay, J. D.; Clore, G. M.; Gronenborn, A. M. *Biochem.* **1992**, *31*, 3442.

(24) Roos, G.; Messens, J.; Loverix, S.; Wyns, L.; Geerlings, P. *J. Phys. Chem. B* **2004**, *108*, 17216.

(25) Joshi, H. V.; Meier, M. S. *J. Am. Chem. Soc.* **1996**, *118*, 12038.

(26) Van Duijnen, P. T.; Thole, B. T.; Hol, W. G. *Biophys. Chem.* **1979**, *9*, 273.

(27) Aqvist, J.; Luecke, H.; Quiocho, F. A.; Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 2026.

(28) Rullmann, J. A.; Bellido, M. N.; van Duijnen, P. T. *J. Mol. Biol.* **1989**, *206*, 101.

(29) (a) Van Duijnen, P. T. *Enzyme* **1986**, *36*, 93. (b) Náray-Szabó, G. *J. Mol. Struct. (THEOCHEM)* **2000**, *500*, 157.

(30) (a) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553. (b) Simon, S.; Duran, M.; Dannenberg, J. J. *J. Chem. Phys.* **1996**, *105*, 11024.

(31) (a) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117. (b) Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *106*, 5151. (c) Cammi, R.; Mennucci, B.; Tomasi, J. *J. Phys. Chem. A* **2000**, *104*, 5631. (d) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2002**, *117*, 43.

(32) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision A.1; Gaussian, Inc.: Pittsburgh, PA, 2003.

(33) Gross, K. C.; Seybold, P. G.; Peralta-Inga, Z.; Murray, J. S.; Politzer, P. *J. Org. Chem.* **2001**, *66*, 6919.

(34) Hashimoto, Y.; Kohri. K.; Kaneko, Y.; Morisaki, H.; Kato, T.; Ikeda, K.; Nakanishi, M. *J. Biol. Chem.* **1998**, *273*, 16544.

(35) Biron, Z.; Khare, S.; Samson, A. O.; Hayek, Y.; Naider, F.; Anglister, J. *Biochemistry* **2002**, *41*, 12687.

(36) van Duijnen, P. T.; Thole, B. T.; Broer, R.; Nieuwpoort, W. C. *Int. J. Quantum Chem.* **1980**, *17*, 651.