Factors That Affect the Stabilization of α-Helices in Short Peptides by a Capping Box

ARTICLE *in* BIOCHEMISTRY · FEBRUARY 1996

Impact Factor: 3.02 · DOI: 10.1021/bi9513766 · Source: PubMed

Factors That Affect the Stabilization of α-Helices in Short Peptides by a Capping Box

Michael Petukhov,[‡,§,||] Noboru Yumoto,[§] Sachiko Murase,[§,⊥] Ritsuko Onmura,[§] and Susumu Yoshikawa*[,§]

Pacific Institute of Bioorganic Chemistry, RAS, Vladivostok 690022, Russia, and Osaka National Research Institute, AIST, Ikeda, Osaka 563, Japan

Received June 19, 1995; Revised Manuscript Received September 19, 1995[⊗]

ABSTRACT: It was reported recently that the capping box sequences of four N-terminal residues are very important for the stabilization of α-helices in proteins and peptides. To elucidate factors that affect the stabilization of α-helices in short peptides by this motif, we analyzed conformational properties of side chains of five N-terminal residues in several analogs of neuropeptide Y (NPY). The analysis revealed three previously unreported factors that appear to be important for stabilization of an α-helix: (a) a second capping box hydrogen bond for the side chains of Ser, Thr, and Cys; (b) long-range electrostatic interactions between the first (N-cap) and fifth (N4) residues; and (c) capping interactions of α-amino groups with the N4 residue. These factors were incorporated into the parameter set of a recently published, statistical mechanics approach that showed excellent accuracy in the prediction of the helical propensities of short peptides in water [Muñoz, V., & Serrano, L. (1995) *J. Mol. Biol. 245*, 275−296, 297−308]. A significant improvement in the agreement between theoretical predictions and experimental data was achieved. The present results also clarify the nature of capping box stabilization of α-helices in peptides and proteins, indicating that the total influence of hydrogen bonding, local interactions between side chains, helix macrodipole−charge/dipole interactions, and solvation possibilities must all be taken into account. All these factors are associated with approximately the same energy, but with different residues at the N-cap position, they may have opposite effects on the helix stability of peptides. Thus, a delicate balance of interactions of different types controls the stabilization properties of N-cap residues in α-helices.

It has been generally accepted that short peptides cannot form stable α-helices in water. However, a large body of experimental information about the helical behavior of peptides has been collected over the past 10 years (Scholtz & Baldwin, 1992). In addition to the importance for the stability of an α-helix of the hydrogen bonding in the backbone and the loss of conformational entropy, it was also recently found experimentally that other factors can play a crucial role in stabilization of the helical conformation of short peptides. Such factors include interactions among side chains (Jimenez et al., 1994; Padmanabhan & Baldwin, 1994), electrostatic interactions among polar and charged groups at termini with a helix macrodipole (Lockhart & Kim, 1992, 1993), and so-called capping interactions at N- and C-termini where polar side chains can form additional hydrogen bonds with free NH− and CO− groups of the main chain (Harper & Rose, 1993). In particular, it was shown experimentally that the choice of the N-cap residue has a major effect on the helical content of the peptides under investigation (Forood et al., 1993; Chakrabartty et al., 1993a; Lyu et al., 1993; Yumoto et al., 1993; Zhou et al., 1994).

Theoretical investigations of the problem of helix−coil transitions in peptides have a long history. The first statistical mechanics approaches appeared more than 30 years ago (Zimm & Bragg, 1959; Lifson & Roig, 1961). The predictions of the theories were in excellent agreement with experimental results for long peptides with repetitive sequences in organic solvents but failed in the case of short peptides with natural sequences in water (Gans et al., 1991). The reason for the discrepancy is the fact that side-chain−side-chain interactions and electrostatic interactions between charged or polar groups with a helix macrodipole, as well as capping interactions, were ignored. After modifications (Vasquez et al., 1987, 1988; Finkelstein et al., 1991), the theories have been significantly improved with respect to their ability to yield qualitative predictions, but quantitative predictions still continue to be a problem. Very promising statistical mechanics approaches to the prediction of helical propensities of short peptides were reported recently (Doig et al., 1994; Muñoz & Serrano, 1994, 1995a,b) and were shown to have excellent accuracy in the prediction of the helical contents of peptides of a very wide range of lengths and sequences.

However, when we used these two approaches to predict the helical propensities of a series of analogs of the C-terminal part of porcine neuropeptide Y (NPY) that include several typical capping box sequences, the results failed to fit the experimental data. The correlation coefficients *R* were only 0.25 and 0.35 for the HELIX (Doig et al., 1994) and AGADIR (Muñoz & Serrano, 1994, 1995a,b) programs, respectively. In the present study, we analyzed conformational properties of N-terminal residues involved in the

* To whom correspondence should be addressed at the Department of Organic Materials, Osaka National Research Institute. Fax: +81-727-51-9628.

‡ Pacific Institute of Bioorganic Chemistry.

§ Osaka National Research Institute.

|| M.P. is a recipient of a fellowship from the Foreign Researcher Invitation Program of the Agency of Industrial Science and Technology (Japan).

⊥ S.M. is a recipient of an STA Fellowship from the Research Development Corp. of Japan.

⊗ Abstract published in *Advance ACS Abstracts,* December 15, 1995.

0006-2960/96/0435-387$12.00/0

capping box stabilization. On the basis of the analysis we attempted to extend Muñoz and Serrano's approach to explain helical propensities obtained from experiments with analogs of the helical part of NPY (Yumoto et al., 1993) and other peptides with capping box sequences (Forood et al., 1993) as well as nonpublished peptides that were synthesized in our laboratory for studies of capping box stabilization.

Our modifications introduced three new factors that appear to be important for stabilization of an α-helix: (a) a second capping box hydrogen bond for the side chains of Ser, Thr, and Cys; (b) long-range electrostatic interactions between N-cap and N4 residues [the nomenclature for the description of positions of amino acids in an α-helix was taken from Richardson and Richardson (1988)]; and (c) capping interactions of α-amino groups with the N4 residue. The theoretical calculations using the modified set of parameters significantly improved the agreement between theoretical predictions and experimental results (R was close 0.9) and revealed further details of the nature of capping box stabilization of α-helices in peptides. We found that calculations based on statistical mechanics can provide reasonable quantitative estimates of the helical contents of the peptides under investigation if energy-related parameters account for all of the most important interactions that take place near the N-termini of the molecules.

MATERIALS AND METHODS

Synthesis of Peptides and CD Measurements. The peptides were synthesized on an automated solid-phase peptide synthesizer (Shimadzu PSSM-8), using Tenta Gel TG-RAM resin and Fmoc chemistry with (benzotriazol-1-yloxy)tris-(pyrrolidino)phosphonium hexafluorophosphate and *N*-hydroxybenzotriazole as coupling reagents. Peptides were cleaved from the resin by using trifluoroacetic acid and purified by reverse-phase HPLC on a C_{18} column. Fmoc-L-amino acids, reagents for peptide synthesis, and Tenta Gel TG-RAM resin were purchased from Shimadzu. The purity of each peptide was assessed by analytical reverse-phase HPLC on a C_{18} column. Molecular masses were confirmed by mass spectrometry on a time-of-flight mass spectrometer (Shimadzu/Kratos Kompact MALDI II) with matrix-assisted laser desorption ionization.

CD[1] measurements were made with a spectropolarimeter (Jasco, Model J-500A, equipped with a data processor, DP-501). CD spectra were recorded at 5 °C in a cell with a 0.2-cm path length that was thermostatically controlled. Each peptide was dissolved in 5 mM 7-*N*-[tris(hydroxymethyl)-methyl]-2-aminomethanesulfonate (TES)−NaOH buffer (pH 7.5) at a concentration of approximately 10 μM. α-Helical contents were calculated using $[\theta]_{222}$ values and the following equation (Scholtz et al., 1991): $f_H = [\theta]_{222}/(-40\,000(1 - 2.5/n))$, where f_H represents the α-helical content and n represents the number of peptide bonds.

Calculations. Calculations based on statistical mechanics were carried out with the AGADIR and HELIX computer programs. The executables and source code of AGADIR were kindly provided by the authors of the program (Muñoz & Serrano, 1994, 1995a,b). The following energy contributions were included in the calculations of the partition functions of all possible helical segments in AGADIR: hydrogen bonding in the backbone; loss of entropy required to put amino acids into helical dihedral angles; $i,i+3$ and $i,i+4$ interactions between side chains; capping interactions at N- and C-termini; and interactions of charged and polar groups with the helix macrodipole. The dependence of the various energy parameters on pH and temperature was taken into account, as implemented in the AGADIR algorithm (Muñoz & Serrano, 1995a,b). The FORTRAN source code of the HELIX program was obtained via anonymous FTP access, as reported by the authors of the program (Doig et al., 1994). The correct installation of both the AGADIR and the HELIX programs was checked by the series of test calculations on the previously published peptide sequences.

Molecular modeling calculations of the lowest energy conformations of amino acids at the N-cap were performed with Quanta 4.0 and CHARMm 2.2 on an INDIGO 2 Silicon Graphics computer. Computer models of the right-handed α-helix of NPY(12-36) analogs substituted with Thr, Asp, and Glu at N-cap were generated with the sequence BUILDER of Quanta 4.0. In all cases, the first 12 residues (Xaa-Pro-Ala-Glu-Asp-Leu-Ala-Arg-Tyr-Tyr-Ser-Ala) were used for modeling. A complete grid search, followed by 50 steps of a steepest descent minimization algorithm, was used to find the lowest energy conformations of Thr, Asp, and Glu residues in N-cap-substituted analogs of NPY(12-36). During the grid search, φ, ψ, and all dihedral angles of the side chains of N-cap as well as all dihedral angles of the side chains of the N3 and the N4 residues were varied from 0° to 360° with grid steps of 60°. Energy calculations were performed for one molecule in a vacuum and included the following contributions: van der Waals forces, hydrogen bonding, electrostatic interactions ($\epsilon = 2r$), torsion barriers, and SHAKE constraints for bond lengths. In order to eliminate the need to renew the interaction list, the complete list of all possible interacting pairs (cutoff = 99) was used in the energy calculations. α-Amino groups were taken in a protonated form. In all cases, the standard CHARMm 2.2 set of parameters was used.

RESULTS

Panels a and b of Figure 1 show the correlations of original HELIX and original AGADIR calculations and the experimentally observed values of helical contents for analogs of NPY(12-36) substituted at the N-terminal residue with almost all natural amino acids (Yumoto et al., 1993). In many cases, the predictions fall outside the range of experimental error, and correlation coefficients were only 0.25 and 0.35, respectively.

In the case of HELIX, the calculations even do not predict the expected shift in helicities of the peptides due to the presence of four aromatic residues in the primary structure. The slope of the correlation graph also was close to zero, indicating the absence of any correlation with relative helical propensities determined by CD.

The calculations with AGADIR gave considerably better correlations with experimental results. The aromatic shift in the CD spectrum was predicted correctly, but in many cases, AGADIR failed to predict the experimentally determined helical propensities of the peptides. In the majority of these cases, the N-cap position was occupied by a residue

[1] Abbreviations: CD, circular dichroism; TFE, trifluoroethanol; θ_{222}, mean residue ellipticity at 222 nm.

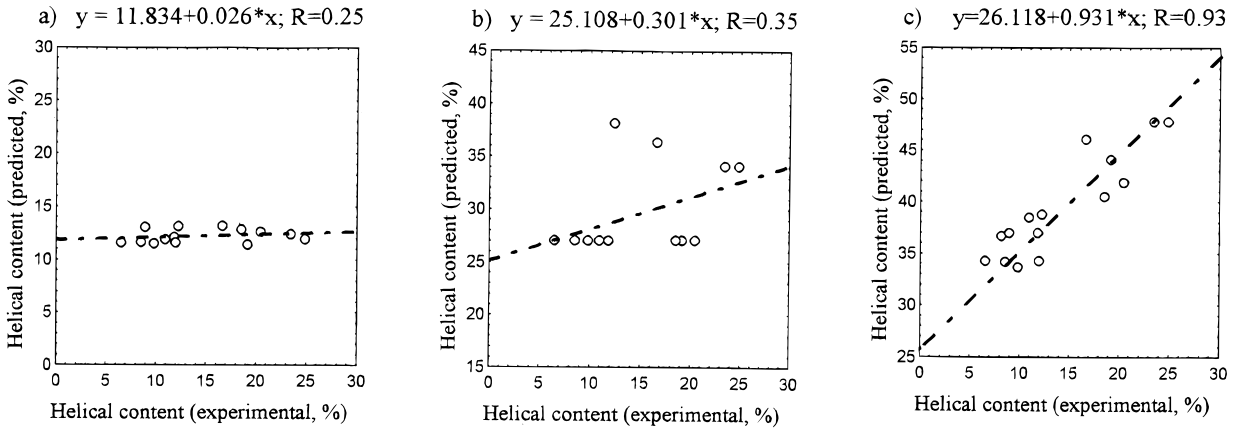

FIGURE 1: Correlation between uncorrected, experimentally determined helicities (Yumoto et al., 1993) and the results of calculations for analogs of NPY(12-36) (a) using the HELIX program, (b) using the original set of AGADIR parameters, (c) using the modified set of AGADIR parameters.

Table 1: Sequences and Results of CD Measurements of the Helical Contents for the Previously Unreported Peptides Used in This Study[a]

peptide (sequence)	$-[\theta]_{222}$ (deg·cm^2/dmol)	helical content: exptl[b] (%)
TA (TPAEAAAAAAYKA-NH_2)	6 800	21.0 (37.2)
TD (TPAEDAAAAYKA-NH_2)	16 700	52.0 (68.5)
DA (DPAEAAAAAYKA-NH_2)	6 500	20.1 (36.3)
DD (DPAEDAAAAYKA-NH_2)	11 600	36.0 (52.4)
S1* (SEDELKAAEAAFKRHGPY-NH_2)	9 381	27.0 (35.7)
A1 (AEDELKAAEAAFKRHGPY-NH_2)	6 958	20.1 (28.8)
S1′ (SPAELKAAEAAFKRHGPY-NH_2)	8 796	25.3 (34.0)
A1′ (APAELKAAEAAFKRHGPY-NH_2)	3 147	9.0 (17.7)
S12′NPY (SEDEDLARYYSALRHYINLITRQRY-NH_2)	7 809	21.4 (46.4)
A12′NPY (AEDEDLARYYSALRHYINLITRQRY-NH_2)	8 592	23.6 (48.6)
ND17 (NAQEDAAAAEAEAKGGY-NH_2)	12 178	35.4
CD17 (CAQEDAAAAEAEAKGGY-NH_2)	8 403	24.4
NE17 (NAQEEAAAAEAEAKGGY-NH_2)	11 976	34.8
CE17 (NAQEEAAAAEAEAKGGY-NH_2)	8 508	24.7

[a] All peptides were analyzed at pH 7.5 in 5 mM TES−NaOH buffer at 5 °C. The percentage of a helicity was calculated by the method of Scholtz et al. (1991). [b] The values corrected for aromatic contributions are given in parentheses. The following corrections of aromatic contributions of Tyr and Phe within the α-helix to $[\theta]_{222}$ were used: TA, DA, TD, and DD peptides, −5000 deg·cm^2/dmol (Chakrabartty et al., 1993b); S1*, A1, S1′, and A1′, −3000 deg·cm^2/dmol (Chakrabartty et al., 1993b); A12′NPY and S12′NPY, −9000 deg·cm^2/dmol (see Discussion). No corrections were made for the terminal Tyr separated from the rest of the molecules by Gly-Gly or Gly-Pro gaps (Chakrabartty et al., 1993b).

that has a side chain that can form a hydrogen bond to the main chain of the α-helix. Since the five residues at the N-terminus of NPY(12-36) (Xaa-Pro-Ala-Glu-Asp) have the so-called capping box sequence that was shown to be important for stabilization of a helix in proteins and peptides (Harper & Rose, 1993; Zhou et al., 1994; Jimenez et al., 1994), we assumed that a main reason for the disagreement with experimental data was poor capping box energy parameters in the original set of parameters. Indeed, Muñoz and Serrano assigned the values on the basis of only few experiments where N-capping propensity of Cys even was not tested (Forood et al., 1993; Zhou et al., 1994). For example, side chains of Glu, Gln, and Cys can form capping box hydrogen bonds, but their capping box free energies were assigned to be +0.25, −0.1, and +0.08 kcal/mol, respectively, in the original set of energy parameters (Muñoz & Serrano, 1995a). In our experimental studies, these amino acids always exhibit significant N-capping potential.

Substitutions by Asn and Asp yield the highest propensities in a very similar N-terminal sequence (Xaa-Pro-Ala-Glu-Ala; Forood et al., 1993) but did not do so in the case of NPY analogs where the highest helical propensities were associated with Thr and Ser (Yumoto et al., 1993). The probable reason for the discrepancy is the long-range electrostatic interactions of N-cap side chains with residues outside the capping box. Poisson−Boltzmann electrostatic calculations (data not presented) revealed the existence of a significant negative electrostatic potential as a result of charged Asp-5 at the site of the N-cap side chain in the capping box conformation. Since, in the work of Forood et al. (1993), the fifth position was occupied by the neutral residues (Ala), we postulated that the difference in N-cap helical propensities can be explained by the presence of Asp-5.

To examine our hypothesis, we synthesized four model alanine-based peptides (peptides TA, TD, DA, and DD in Table 1). The sequences were designed to exclude any complicated interactions within the helix to simplify the interpretation of the results of CD measurements, which are also presented in Table 1. As expected, the presence of a negative charge at N4 changed the N-cap helical propensities for Thr and Asp, indicating the importance of electrostatic interactions between side chains of N-cap and N4.

To elucidate the nature of the capping box stabilization and to obtain a more reasonable set of parameters for N-capping, we attempted to separate the contributions to the stabilization of the capping box of several physical factors that have been shown experimentally to have a major effect

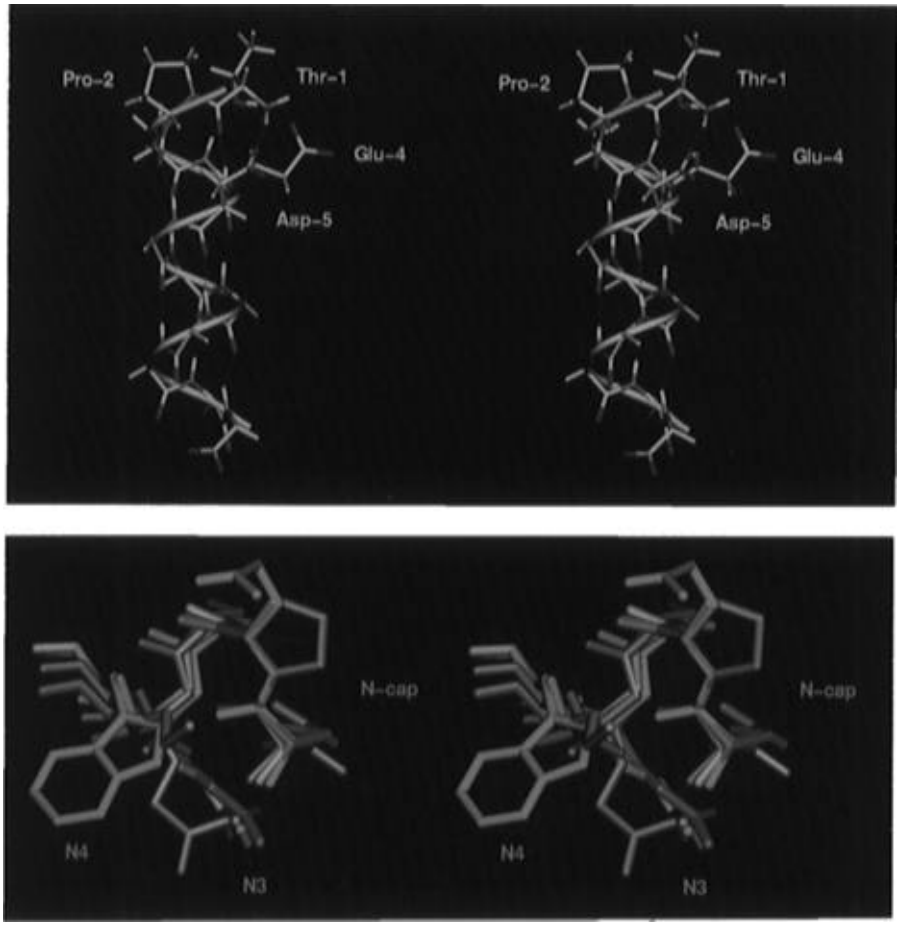

FIGURE 2: Stereo pair of photographs showing (a, top) the capping box conformation of the N-cap residue (Thr) at the N-terminus of NPY(12-36) (the most important capping box contacts are shown by dashed lines) and (b, bottom) superposition of the capping box confromations taken from crystal structures of proteins from the Brookhaven Protein Data Bank (Bernstein et al., 1977): 1BP2 (Asn-Ala-Cys-Glu-Ala), shown in pink; 1MBO (Ser-Glu-Cys-Glu-Trp), shown in yellow; 2LHB (Ser-Ala-Ala-Glu-Lys), shown in green; and the lowest energy conformation obtained for Thr-NPY(12-36) (Thr-Pro-Ala-Glu-Asp) as described in Materials and Methods, shown in blue.

on the α-helix stability of short peptides in water. However, since natural amino acids have different conformational possibilities at the N-terminus, we used the Quanta 4.0 molecular modeling program to find the lowest energy conformations for several typical amino acids at the N-cap of NPY(12-36) (see Materials and Methods).

The typical low-energy conformation of the capping box in NPY(12-36) is presented in Figure 2a. In this conformation, the side chain of Thr (and, therefore, those of Ser and Cys) can form two hydrogen bonds, namely, $O_\gamma(S_\gamma)$ with the amide proton of Glu at N3 and, simultaneously, HO_γ-(HS_γ) with the carboxyl group of Glu at N3, while Asn, Asp, Gln, and Glu can form one only, and the others can form none. This observation can explain why, in spite of the absence of any additional stabilizing interactions, Thr, Ser, and Cys are often associated with some of the highest helical propensities at N-cap positions of α-helices in proteins and peptides (Serrano & Fersht, 1989; Forood et al., 1993; Lyu et al., 1993; Yumoto et al., 1993). Another observation is that, in the lowest energy capping box conformation, the α-amino groups of N-terminal residues are located very close to both the carboxyl groups of Glu/N3 and Asp/N4, with a good possibility for hydrogen bonding. The lowest energy conformations obtained for other substitutions at N-cap (Asp and Glu) reveal the very close arrangement of side-chain to side-chain and side-chain to main-chain contacts, which appear to play a role in stabilization of the α-helix by the capping box.

The lowest energy conformations of the capping box obtained in these calculations were very similar to those found in proteins (Seale et al., 1994) in terms of dihedral angles and relative disposition of side chains. This fact is clearly demonstrated by the superposition of the several capping box conformations presented in Figure 2b. Although the reported pattern of hydrogen bonds in the capping box (the side chain of N-cap to the amide of N3 and the side chain of N3 to the amide of N-cap) was possible, we found also a more plausible pattern of hydrogen bonds in particular sequences of the analogs of NPY. This pattern gains one hydrogen bond as compared with a regular one without any distortion of the typical capping box conformation (see Figure 2).

Modification of the Energy Parameters of the Capping Box. There are many experimental and theoretical results related to the importance for conformational stability of different physical interactions at N-termini of α-helices. In this analysis, we assumed that the following interactions are the most important: hydrogen bonding of side chains and α-amino groups at the N-cap; side-chain−side-chain interactions between N-cap and N2 and between N-cap and N3; long-range electrostatic interactions between the side chains at N-cap and N4; dipole/charge and dipole/dipole interactions between the helix macrodipole and charged or polar groups

Table 2: Main Contributions to the Free Energy of Capping Box Stabilization in the Analogs of NPY(12-36) (Xaa-Pro-Ala-Glu-Asp...) and Results of CD Measurements

N-cap	H-bonds (kcal/mol)	contribution of α-amino group (kcal/mol)	N-cap/N4 electrostatic interactions (kcal/mol)	N-cap charge/helix macrodipole interactions[b] (kcal/mol)	N-cap dipole/helix macrodipole interactions (kcal/mol)	N-cap/N2 side-chain interactions (kcal/mol)	N-cap/N3 side-chain interactions (kcal/mol)	total (kcal/mol)	helical content: exptl[d] (%)
Ala		−0.44	0.00	0.00	0.00	0.0	0.0	−0.44	8.5 ± 1.0
Gly		−0.43						−1.44*c*	20.4 ± 2.0
Ser	2 (−0.775)	−0.56	0.00	0.00	0.00	0.0	0.0	−2.11	23.4 ± 2.0
Thr	2 (−0.775)	−0.56	0.00	0.00	0.00	0.0	0.0	−2.11	24.8 ± 2.0
Asn	−0.775	−0.67	0.00	0.00	−0.30	0.0	0.0	−1.73	16.6 ± 1.0
Asp^-	−0.775	−0.59	+0.30	−0.34	0.00	0.0	+0.20	−1.19	12.2 ± 1.0
Gln	−0.775	−0.54	0.00	0.00	−0.30	0.0	0.0	−1.59	19.1 ± 2.0
Glu^-	−0.775	−0.46	+0.30	−0.26	0.00	0.0	0.0	−1.18	10.9 ± 1.0
His^+		−0.54	−0.20	+0.33	0.00	0.0	−0.20	−0.61	8.1 ± 1.0
Lys^+		−0.59	−0.20	+0.43	0.00	0.0	−0.10	−0.41	9.8 ± 1.0
Arg^+		−0.54	−0.20	+0.29	0.00	0.0	−0.10	−0.60	6.5 ± 1.0
Phe		−0.56	0.00	0.00	0.00	0.0	−0.40	−0.96	11.8 ± 1.0
Tyr		−0.56	0.00	0.00	0.00	0.0	−0.40	−0.96	8.9 ± 1.0
Trp		−0.51	0.00	0.00	0.00	0.0	−0.40	−0.91	8.4 ± 1.0
Val		−0.47	0.00	0.00	0.00	0.0	0.0	−0.47	11.9 ± 1.0
Cys	1.24 (−0.775)[a]	−0.40	0.00	0.76 (−0.3)[a]	0.00	0.0	0.00	−1.56	18.5 ± 2.0

[a] These estimates were obtained by taking into consideration partial ionization of the side chain of Cys. [b] The estimates of the electrostatic interactions of charged amino acids at N-cap with the helix macrodipole were taken from Muñoz and Serrano (1995a). [c] Gly is more stabilizing than Ala by 0.5−1.0 kcal/mol at N-cap in barnase (Serrano & Fersht, 1989). Therefore, we assigned a difference of 1.0 kcal/mol in comparing Ala to Gly in the capping box. [d] The experimental values were taken from Yumoto et al. (1993) without corrections for the aromatic contribution to the CD spectra (see Discussion). Estimates of experimental error are given after the (±) mark.

at the N-terminus; and capping interactions of the α-amino group with the side chains of the N3 and N4 residues.

Hydrogen bonding in the capping box is thought to be a major factor in helix stabilization. We took the estimated value of −0.775 kcal/mol for the contribution of each hydrogen bond from the N-cap side chain to the stabilization free energy of the capping box (Muñoz & Serrano, 1995a).

We are unaware of any experimental data for electrostatic interactions between N-cap and N4. However, since these groups are not in direct contact, the contribution must be quite small and probably does not exceed the values obtained for the *i*,*i*+4 Glu/Lys ion pair (−0.5 kcal/mol; Gans et al., 1991). Therefore, we assigned a value of +0.3 kcal/mol to the Asp/Asp and Asp/Glu pairs and a value of −0.2 kcal/mol to the Asp/Arg, Asp/Lys, and Asp/His pairs.

It is well-known that interactions between the helix macrodipole and charged and polar groups at the N-terminus are also very important for stability of the α-helix. However, it is difficult to calculate their energies correctly because these interactions occur in the complicated heterogeneous environment of the border between the peptide and water. In this analysis, we used experimental estimates obtained from observed shifts in several charged and polar groups at the N-terminus of the α-helix (Lockhart & Kim, 1992, 1993). In particular, it has been shown that polar 4-(methylamino)-benzoil (MABA) destabilizes the α-helix in alanine-based peptides by +0.57 kcal/mol. Lockhart and Kim also found, by NMR, that the conformation of MABA at the terminus of a helix was fixed, with an angle of 20° between the dipole moment of MABA (8 D) and the axis of the helix. Therefore, it is easy to estimate the energy of dipole/dipole interactions per unit of the MABA dipole moment along the helical axis: 0.57/(cos (20)·8D) = 0.075 kcal/(mol·D). Then the stabilization dipole/dipole energy for Asn and Gln side chains in the capping box conformation will be close to −0.3 kcal/mol (4.1 D × 0.075), where 4.1 D is the value of the dipole moment of the amide groups of Asn and Gln, as calculated using the CHARMm 2.2 set of charges for peptides.

The interactions of the charged side chains at N-cap with the helix macrodipole are another important contribution to the stabilization of the capping box. In the last version of the AGADIR algorithm, these interactions included in a special term that is added automatically to the total free energy of the particular helical segment (Muñoz & Serrano, 1995a). Standard values for these interactions are listed in the Table 2.

It is also important to note that the potential from the helix macrodipole near the termini of the helix depends very weakly on the length of the helix because of the screening by water of long-range electrostatic interactions. It has been shown that the electrostatic interactions inside the first two turns of the helix should account for almost 90% of the interaction energy between the macrodipole of the helix and the charged or polar group near the terminus of the α-helix (Hol et al., 1978; Aqvist et al., 1991).

The α-amino group of the N-cap residue can also affect the helical content of a peptide under two sets of conditions, at least: (a) if it is protonated via electrostatic interactions with the helix macrodipole and with charged side chains of the nearest amino acids, if any; and (b) if it is uncharged through hydrogen bonding with the side chains of N3 or N4, residues, if these residues can accept the amide proton. Therefore, ionization of the N-terminal amino group should be also important for the stabilization of the helix. It was shown experimentally that pK_a values of α-amino groups in alanine-based peptides depend on the nature of the N-cap amino acids (Forood et al., 1993; Chakrabartty et al., 1993a) with a shift of 1.5 pH units approximately below the standard values measured for the free amino acids (Cantor & Schimmel, 1980). Since all experiments with analogs of NPY were performed at pH 7.5, and the expected pK_a constants of the α-amino groups and the side chain of Cys are close to 7.5, we calculated the ionization degree *I* of the

groups from the classical equations:

$$I_{\text{basic}} = 1/(1 + (10^{pK_a}/10^{pH}))$$

$$I_{\text{acidic}} = 1/(1 + (10^{pH}/10^{pK_a}))$$

The values of pK_a of the α-amino groups were taken from Chakrabartty et al. (1993a) as follows: Gln, 7.72; Val, 8.14; Ile, 8.24; Ala, 8.35; Met, 7.83; Pro, 8.85; Leu, 8.31; Thr, 7.62; Gly, 8.51; Ser, 7.63; and Asn, 7.07. The values of the pK_a for the α-amino groups of other amino acids, as well as for side chains of Cys, were obtained by shifting the pK_a by 1.5 units below the standard values (Cantor & Schimmel, 1980).

As we can see from Figure 2a, in the capping box conformation of NPY(12-36), the α-amino group is located very close to both side chains of N3 and N4, which are negatively charged. Therefore, if the α-amino group is protonated, we have to incorporate into the capping box energy the favorable interactions with the N3 and N4 side chains and the unfavorable interactions with the helix macrodipole. In the opposite case, if the α-amino group is neutral, we add one hydrogen bond to the capping box stabilization energy. Since the energy cost of one ion pair equals approximately −0.5 kcal/mol (Gans et al., 1991) and the energy of the interaction with the helix macrodipole has a value of +0.6 kcal/mol (Muñoz & Serrano, 1995a), we can assume that the total electrostatic contribution of a protonated α-amino group is 2(−0.5) + 0.6 = −0.4 kcal/mol.

Interactions between side chains in peptides are complicated by the mixed influence of hydrogen bonding, the contribution of lost entropy, electrostatic interactions, and hydrophobic interactions. Since the side chains at N2 and N3 are in direct contact with the N-cap side chain and their relative disposition in the capping box conformation was similar to that of *i*,*i*+4 and *i*,*i*+3 side-chain interactions within an α-helix, we assigned, as a first approximation, *i*,*i*+4 energies from Muñoz and Serrano's set of parameters to the N-cap/N2 pair and *i*,*i*+3 energies to the N-cap/N3 pair that should include all the above-mentioned energy terms (Muñoz & Serrano, 1995a).

It has been shown that strong interactions in the capping box can suppress the N-terminal fraying of a short peptide (Zhou et al., 1994), stabilizing close contacts between N-cap and the side chains of N3 and N4 and, therefore, destabilizing the possible *i*,*i*+4 and *i*,*i*+3 interactions for these side chains within a particular helical segment. Therefore, we did not add these interactions to the total free energy of helical segments when the formation of the hydrogen bonds in a capping box was possible.

Table 2 presents our estimates for the capping box stabilization energies for the analogs of NPY(12-36) and the experimentally determined helicities, as measured by CD (Yumoto et al., 1993). The correlation between the modified energies of capping box stabilization and the experimentally determined helicities was very high ($R = 0.91$).

It is of interest also to examine how this analysis can be applied to other peptides that have different capping box sequences. One can see from Figure 2a that the lowest energy conformation of the capping box presupposes no restrictions on the side chains at N1 and N2 when the maximum number of hydrogen bonds is saturated. Therefore, this conformation must be very frequent for any capping box-like sequence. Moreover, very similar conformations were found recently in crystal structures of proteins (Seale et al., 1994). It is of interest that, despite the very different capping box sequences listed in Table 1 of the report by Seale et al. (1994), all but one (5CYT) of their structures were very close to that presented in Figure 2a with respect to the arrangement of capping box contacts (RMSD of the atoms involved in the contacts was close to 0.2 Å). However, in the proteins, the conformations are almost exclusively stabilized by contacts between the hydrophobic side chains of N′ and N4. It was very recently reported that such a structural motif (hydrophobic interactions between N′ and N4) also plays an important role in stabilization of the α-helix in peptides (Muñoz et al., 1995). In the case of derivatives of NPY, different interactions between the polar side chain at N4 and the α-amino group of N-cap occur in the similar conformation. Figure 2b shows the superposition of the three crystal structures of the capping boxes along with the lowest energy conformation for the analog of NPY-(12-36). We presented in the superposition these three crystal structures for clarity of the picture only. The conformations of other crystal structures, listed in the paper of Seale et al. (1994), were very close to those presented in the superposition.

A similar analysis was performed for the analogs of NPY-(13-36). In this case, formation of a capping box was also possible, but since the sequence of the first five N-terminal amino acids was somewhat different from that of NPY(12-36), namely, Xaa-Ala-Glu-Asp-Leu, the capping box stabilization energies appeared to differ from those of NPY(12-36) also. First, a shift in one amino acid drastically changed the side-chain−side-chain contacts in the capping box. Next, there was no charged amino acid at the N4 position, and therefore, N-cap/N4 electrostatic interactions always made a contribution of zero. Additionally, if N-cap is occupied by Gly, no additional solvation of free NH− groups near the N-terminus should exist because of the influence of the bulky side chain of Glu at N2 (el Masry & Fersht, 1994). Finally, the α-amino group also made a near-zero contribution to the capping box stabilization as follows: (a) in the case of protonated α-amino groups, unfavorable interactions with the helix macrodipole were balanced by interactions with the negatively charged side chain of Asp at N3; and (b) in the case of a neutral α-amino group, there was no good acceptor for hydrogen bonding since N4 was occupied by Leu and the side chain of Asp at N3 was not long enough for good hydrogen-bonding geometry. Table 3 summarizes the energy contributions to the capping box stabilization in NPY(13-36).

Several peptides with capping box sequences and also with a so-called "hydrophobic-staple" motif (Seale et al., 1994; Muñoz et al., 1995) at N-termini (Ac-YM-Xaa-EDEL-KAAEAAFKRHGPT-NH_2) were analyzed by Zhou et al. (1994). However, these peptides were excluded from the set of control peptides because the correct parameters for these peptides cannot be obtained without incorporating of energy contributions from the hydrophobic interactions between side chains of N′ and N4. Unfortunately, such estimates have not been reported yet.

Table 3: Main Contributions to the Free Energy of Capping Box Stabilization in the Analogs of NPY(13-36) (Xaa-Ala-Glu-Asp-Leu...) and Results of CD Measurements

N-cap	H-bonds (kcal/mol)	contribution of α-amino group (kcal/mol)	N-cap charge/helix macrodipole interactions[a] (kcal/mol)	N-cap dipole/ helix macrodipole interactions (kcal/mol)	N-cap/N2 side chain interactions (kcal/mol)	N-cap/N3 side chain interactions (kcal/mol)	total (kcal/mol)	helical content: exptl[b] (%)
Ala		+0.08	0.00	0.00	0.00	0.00	+0.08	13.7 ± 1.4
Ser	2 (−0.775)	+0.06	0.00	0.00	0.00	0.00	−1.49	19.3 ± 2.0
Thr	2 (−0.775)	+0.06	0.00	0.00	0.00	0.00	−1.49	19.3 ± 2.0
Asn	−0.775	+0.03	0.00	−0.30	0.00	0.00	−1.05	14.1 ± 1.4
Asp^-	−0.775	+0.05	−0.34	0.00	+0.10	+0.10	−0.87	12.3 ± 1.2
Gln	−0.775	+0.06	0.00	−0.30	−0.25	0.00	−1.27	13.1 ± 1.3
Glu^-	−0.775	+0.08	−0.26	0.00	+0.20	+0.20	−0.56	15.8 ± 1.6
Lys^+		+0.06	+0.43	0.00	−0.30	+0.10	+0.29	11.1 ± 1.1
Arg^+		+0.05	+0.29	0.00	−0.30	+0.10	+0.14	11.3 ± 1.1
Pro		+0.09	0.00	0.00	0.00	0.00	+0.09	12.3 ± 1.2

[a] See footnote *b* in Table 2. [b] The experimental values were taken from Yumoto et al (1993) without corrections for the aromatic contribution to the CD spectra (see Discussion). Estimates of experimental error are given after the (±) mark.

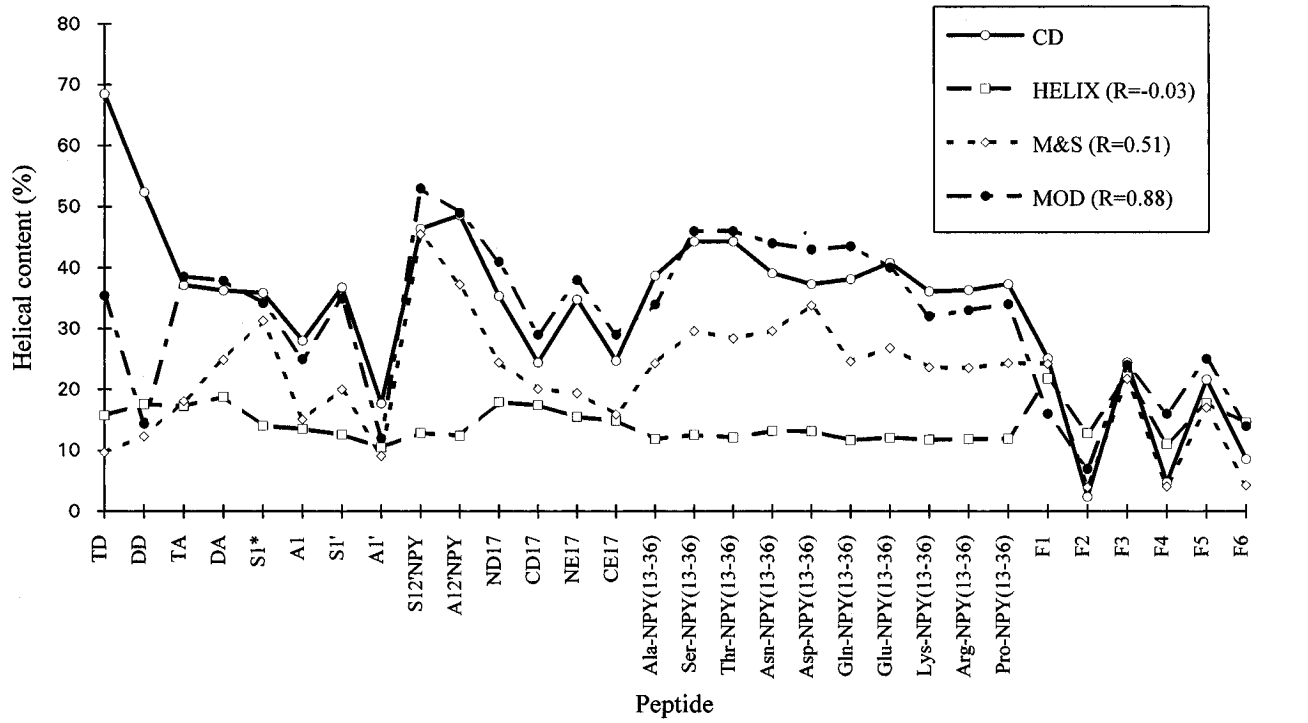

FIGURE 3: Correlation between corrected, experimentally determined helicities and the results of calculations for previously reported peptides (Forood et al., 1993; Yumoto et al., 1993) that are not related to the capping box sequence of NPY(12-36) peptides and for other peptides whose sequences and experimentally measured helicities are listed in Table 1. The CD line presents the experimental helical contents as determined by CD; the HELIX line shows the results of calculations of the HELIX program; the M&S line shows the results of AGADIR calculations with the original set of parameters; and the MOD line shows the results of calculations using the modified set of AGADIR parameters.

The helicities of several alanine-based peptides with Xaa-Pro-Ala-Glu-Ala sequences were examined by Forood et al. (1993). In our study, we also measured the helicity of two peptides (TA and DA in Table 1) with different primary structures but with the same capping box sequence. The presence of uncharged Ala at N4 affects two contributions to the stabilization of the capping box: (a) N-cap/N4 electrostatic interactions and (b) the contribution of the α-amino group. However, since, in the study of Forood et al. (1993), the measurements were made at pH 7.0, the ionization of the α-amino groups will differ from that of the case of TA and DA peptides. Taking into account these considerations, we obtained values for the stabilization of the capping box (data not presented) for peptides F1, F2, F3, F4, F5, and F6 as well as for several previously unreported peptides (see Table 1) that had been synthesized in our laboratory for studies of the stabilization of the capping box.

We used original and modified sets of parameters for calculations of the helical contents of peptides of known helicity that contained capping box sequences near their N-termini (Forood et al., 1993; Yumoto et al., 1993), as well as other peptides whose sequences and experimentally measured helicities are presented in Table 1. The correlation plot obtained by using the uncorrected values of the experimentally determined helicities of the analogs of NPY-(12-36) and the results of the calculations using the AGADIR program with a modified set of parameters are shown in Figure 1c. Figure 3 shows the correlation plot for other peptides, taken from the literature or synthesized in our laboratory.

DISCUSSION

Effects of Aromatic Contributions on CD. Far-UV CD spectra of proteins are usually associated with peptide bonds while near-UV CD bands mainly arise from side chains of aromatic residues and S−S bridges (Strickland, 1974). However, it was shown recently (Chakrabartty et al., 1993b) that aromatic residues also significantly affect the CD spectra in the range of 222 nm. Since the sequences of NPY fragments and of some other peptides used in these calculations contain aromatic residues in the middle of the chains, it is expected that using the standard formula $-f_H = [\theta]_{222}/40\,000(1 - 2.5/n)$ for the estimation of the helical contents should, in such cases, significantly underestimate the real values. Therefore, a correction of experimentally determined helicities for the aromatic contribution is required.

The Lifson−Roig approach, as implemented in the HELIX program, predicts the same or even lower levels of helicity for the aromatic residue-containing peptides, as compared with the uncorrected experimental data (Figure 1a). The correlation between the relative predicted and experimentally determined helical propensities was very poor.

The predictions of AGADIR, using the original set of parameters (Figure 1b), gave as much as twice the helicity determined from uncorrected experimental data, in agreement with the expected influence of four Tyr residues on the CD spectrum. However, the correlation between relative helical propensities was again very poor.

Both the general level and the relative propensities of analogs of NPY(12-36) were predicted well by AGADIR using the modified set of parameters (Figure 1c). Moreover, the differences between predicted and uncorrected experimentally determined helicities were approximately constant. The average difference was close to 25% with a standard deviation of 2.6%. No correlation between the differences and helical propensities was observed. Thus, the theory predicts that the four Tyr residues in analogs of NPY(12-36) decrease the observed helicities of these peptides by a constant value of approximately 25%, which corresponds to 9000 deg·cm^2/dmol in θ_{222}.

Unfortunately, the effects on the CD spectrum of multiple aromatic residues in the middle of an α-helix have not yet been investigated. However, it was shown that one Tyr residue contributes approximately 5000 deg·cm^2/dmol to θ_{222} (Chakrabartty et al., 1993b). The effect of two Tyr residues located near N- and C-termini was approximately the same, indicating that the effect can be nonadditive (Chakrabartty et al., 1993b). This conclusion is in general agreement with the predictions of the theory when the modified set of parameters is used.

Studies of the effect of aromatic contributions to CD at different concentrations of TFE revealed a nonlinear dependence at high concentrations of TFE (Chakrabartty et al., 1993b). It is still unclear whether this phenomenon originates from higher helicity, the presence of TFE, or both. However, at low and at moderate concentrations of TFE the effect showed an approximately constant contribution to CD [see Figure 3 of Chakrabartty et al. (1993b)], indicating that the populations of the side-chain conformers of aromatic residues were not distorted significantly by the stabilization of the helix. The change in these populations was shown to have a major effect on CD spectra of peptides (Strickland, 1974). Since the helical contents of the analogs of NPY and those reported by Chakrabartty et al. (1993b) are in a similar range of the helicity, there seems to be no dependence on the level of helicity of an aromatic contribution in the analogs of NPY.

Measurement of the mean residue ellipticity of an Ala-substituted analog of NPY(12-36) at a concentration of TFE of 50% gave a θ_{222} of −29 000 deg·cm^2/dmol, indicating a shift of 7000 ± 2900 deg·cm^2/dmol as compared with the value of −36 000 deg·cm^2/dmol for the expected ellipticity at 100% helicity, calculated from the standard equation. This result is also in excellent agreement with the predictions of the theory.

Given the above-mentioned observations, we corrected the CD measurements of all NPY-related peptides used in this study by adding −9000 deg·cm^2/dmol to the mean residue ellipticity measured by CD. For some other peptides used in our calculations that include single aromatic residues in the middle of helix, we followed the established procedure to correct for the aromatic contribution to CD (Chakrabartty et al., 1993b). For those peptides that include terminal aromatic residues separated from the rest of the molecules by Gly-Gly or Gly-Pro gaps no correction was added (Chakrabartty et al., 1993b).

Another possible way to compare the experimental data with theoretical predictions is to use relative values, taking one of the experimental values as a basis. The main disadvantage of this method is the as much as 2-fold increase in experimental error due to subtraction of two experimental values, each of which incorporates errors. Additionally, in this case, the basis value in each series of peptides is lost from the comparison since it always has a relative value of zero. Taking these defects into account, we can see that using corrected values provides a more adequate method of comparison.

NPY(11-36) Analogs. The HELIX and the AGADIR with both sets of parameters predict that no substitutions of the N-terminal residue affect the helical stability of NPY(11-36) analogs (Xaa-APAEDLARYYSALRHYINLITRQRY-NH_2) and that the helical content of such peptides varies within 1% (data not shown). The predictions are in excellent agreement with experimental data (Yumoto et al., 1993).

It is well-known from conformational analysis that an amino acid residue with a side chain cannot assume helical dihedral angles if Pro is the next residue in the sequence because of steric restrictions, due to side-chain−side-chain interactions (Cantor & Schimmel, 1980). Therefore, Ala-2 of NPY(11-36) is a strong helix-breaker. Since two residues are insufficient for formation of a new stable helical segment, the total helical content of the peptide will more likely depend on the second residue, which should be involved in N-capping interactions. The uncorrected experimental values for the helical contents of NPY(11-36) analogs are close to 8%. An equal value was determined experimentally for the Ala-substituted analog of NPY(12-36) (Yumoto et al., 1993), indicating that this interpretation is not far from reality.

Moreover, if analogs of NPY(11-36) have equal helical content, we can estimate the errors in the experimental procedure. The average uncorrected value of helical content for eight analogs was 8.06%, with an RMSD of 0.63%

(Yumoto et al., 1993). Thus, the experimental error is probably close to 10%. However, in one of the experiments with Ser-NPY(11-36) the experimental value for the helical content deviated from the average by 1.2%, indicating a 16% error.

NPY(12-36) Analogs. Unlike HELIX, AGADIR predicts a strong dependence of the helicity of NPY(12-36) analogs on the first residue (in agreement with the experimental data), but in many cases, the predictions of the theory made with the original set of parameters fall outside the range of experimental error. In particular, the calculations significantly underestimate the helical propensities for almost all amino acids that can participate in hydrogen bonding of the capping box type. The correlation coefficient R between experimental values and theoretical predictions was 0.35. The most significant disagreements between experimental data and theoretical predictions were observed for the following N-cap residues: Ser, 48.4% *vs* 34.1%; Thr, 49.8% *vs* 34.1%; Cys, 43.5% *vs* 27.1%; Gln, 44.1% *vs* 27.1%; and Gly, 45.4% *vs* 27.1%. Hereafter, we use the experimental values corrected for the aromatic contribution for all analogs of NPY, as described above.

When based on the modified set of parameters, calculations using statistical mechanics eliminate the most significant disagreements between experimental and theoretical helicities, yielding the following: Ser, 48.4% *vs* 47.8%; Thr, 49.8% *vs* 47.8%; Cys, 43.5% *vs* 40.5%; Gln, 44.1% *vs* 44.0%; and Gly, 45.4% *vs* 41.9%. The correlation with experimental data is significantly higher than in the case of the original set of parameters ($R = 0.93$).

The nature of the N-capping stabilization by Gly is probably connected to the additional solvation of free NH− groups at the N-termini of helices if the side chain at N2 is not so bulky as to prevent it (Serrano & Fersht, 1989; el Masry & Fersht, 1994). In this study, we took a value of −1.0 kcal/mol for Gly *vs* Ala stabilization from experiments on barnase (Serrano & Fersht, 1989). However, the conditions of solvation of N-termini in short peptides should be different from those in globular proteins, and probably, a more realistic value of Gly *vs* Ala N-capping stabilization energy would be about −1.2 kcal/mol, which eliminates the minor underestimation of the helical propensity of the Gly-substituted analog of NPY(12-36).

NPY(13-36) Analogs. Calculations using HELIX also gave very poor correlations with experimental data for these analogs of NPY. Neither the general level of helicity nor the relative helical propensities were predicted correctly.

Calculations using AGADIR with the original set of parameters gave approximately the same level of agreement with the experimental data ($R = 0.42$). The highest helical propensities were not predicted correctly. In particular, the helical propensities for Ser, 44.3% *vs* 29.6%, Thr, 44.3% *vs* 28.4%, Asn, 39.1% *vs* 29.6%, Gln, 38.1% *vs* 24.5%, and Glu, 40.8% *vs* 26.8%, substituted analogs were strongly underestimated. Use of the modified set of parameters significantly increased the agreement with experimental results ($R = 0.89$), yielding the following values: Ser, 44.3% *vs* 46.0%; Thr, 44.3% *vs* 46.0%; Asn, 39.1% *vs* 43.9%; Gln, 38.1% *vs* 43.5%; and Glu, 40.8% *vs* 40.0%.

Peptides That Are Not Related to NPY. Calculations using HELIX showed excellent accuracy in predicting the helical propensities of peptides described by Forood et al. (1993). The general level and main helical propensities of the peptides were predicted correctly. The correlation coefficient was 0.93. Predictions of AGADIR using the original set of parameters also showed excellent accuracy for these peptides. All predictions were within the limits of experimental error, and the correlation coefficient R was very high (0.96). However, this result is not surprising because these data were used in the refinement of the empirical set of parameters for the theory. The modified set of parameters also successfully predicted the helicities of these peptides. General levels of helicity were somewhat overestimated, and the correlation ($R = 0.88$) was somewhat lower than in the case of the original set of parameters, but all predictions were still close to the range of experimental error.

The two pairs of peptides (TA,DA) and (TD,DD) were constructed specifically to examine the existence of N-cap−N4 electrostatic interactions. However, unexpectedly high levels of helicity of TD and DD peptides were observed (see Table 1). Calculations using HELIX were unable to predict either the general level of helicity or the relative helical propensities. It was predicted that highest and lowest levels of helicity would be associated with the DA and TD peptides, respectively, while exactly the opposite was observed experimentally.

The AGADIR calculations using the original set of parameters also failed to predict correctly the general level of helicity of these peptides as well as their helical propensities because N-cap−N4 electrostatic interactions were ignored. The calculations using the modified set of parameters correctly predicted the general level of helicity of the TA,DA pair of peptides and the relative helical propensities of both pairs of peptides.

Unfortunately, the theory using both sets of parameters was unable to predict correctly the general level of helicity for TD and DD peptides, indicating the existence of an important stabilizing factor that involves the N4 position, if this position is occupied by Asp. We identified one reason that could explain this effect: it is the difference of 0.6 kcal/mol in the "intrinsic" helical propensities (E_{int}) between Asp and Ala in the AGADIR parameter set (Muñoz & Serrano, 1995a). As defined by Muñoz and Serrano (1995a), this term reflects the free energy of assigning the amino acids to helical dihedral angles and it does not depend on whether a residue is located in the middle of a helical segment or close to its termini. Therefore, E_{int} should include two sources of the lost entropy—the backbone and the side chains. Indeed, Gly has no side chain at all, but its E_{int} is one of the highest among natural amino acids [see Table 2 in Muñoz and Serrano (1995a)] because of the high flexibility of the backbone of the residue. Since other natural residues with side chains (except Pro) are known to have similar levels of flexibility in the backbone, as reflected by the similarities in their (ϕ, ψ) maps (Cantor & Schimmel, 1980), the differences observed between their values of E_{int} should mainly reflect the restrictions of rotational freedom of the side chains within the α-helix. However, side chains of terminal residues have fewer restrictions to their rotation as a consequence of the absence of previous or next helix turns. Therefore, their values of E_{int} must be different from those of central residues in the helix. This basic factor makes the "intrinsic helical propensities" of amino acids with side chains dependent on the particular helical segment under consideration. Since the Asp at N4 is disposed in the first turn of helices in TD and DD peptides, its intrinsic propensity

seems to be always close to that of Ala. This effect should be especially important for short peptides, in which the contribution of terminal residues is so great.

The increase of helical contents for substitutions of Ala by Asp at several terminal and central positions of Ala-based peptides was reported by Huyghues-Despointes et al. (1993). The U- and L-shaped dependencies of the helical content on the position of the substitutions by neutral and charged Asp were interpreted in terms of electrostatic interactions with the helix macrodipole and fraying of the helix ends. However, the terminal fraying should decrease electrostatic interactions with the helix macrodipole, because the charge−helix macrodipole interactions require close disposition of charged groups to the helix (Lockhard & Kim, 1992, 1993) and also because the contribution from the first helix turn, which must be absent here, accounts for the main part of the interaction energy between the macrodipole and the charge (Hol et al., 1978; Aqvist et al., 1991). Therefore, it remains unclear why the residues of charged Asp^- at the terminal positions interact with the helix macrodipole as shown in Figure 1b of the paper by Huyghues-Despointes et al. (1993). These experimental data can be also interpreted in terms of fewer restrictions of rotational freedom of the side chain of Asp at central and terminal positions of the α-helices. Indeed, for neutral Asp, no interactions with the macrodipole should exist, and therefore the U-shaped dependence for helical content on the position of Asp is quite expected due to the increase of intrinsic propensities of Asp at the termini of the helix. The interaction of charged Asp^- with the macrodipole reinforces the effect if it occurs near the N-termini, while at the C-termini the factors are opposite to each other. All these drive the dependence of helical content on the position of Asp^- to the L-shaped curve. Thus the experimental results observed by Huyghues-Despointes et al. tend to support our hypothesis. The detailed estimates of the loss of conformational entropy for natural amino acids at terminal and central positions of the α-helix will be published elsewhere.

The HELIX and the AGADIR calculations using the original set of parameters predicted only minor differences in helical contents for the XD17 and the XE17 series of peptides, while significant differences between helical propensities were observed. The correlation was quite poor. The use of the modified approach significantly improved the correlation with experimental data ($R = 0.89$). The helical propensities were predicted correctly. However, the general levels of helicity were slightly overestimated.

The use of both sets of parameters in AGADIR yielded good accuracy in predictions of helical properties of S1*, A1, S1, A1′, S12′NPY, and A12′NPY peptides. The general levels, as well as helical propensities, were predicted correctly. The predictions of relative values were close to the range of experimental error. The correlation coefficients between the experimental data and the results of the calculations for this set of peptides, obtained with both sets of parameters, were close to 0.9. The predictions of HELIX for this set of peptides showed a very poor correlation with the experimental data in terms of general levels of helicity, as well as of relative helical propensities.

Helix Stop Signals. It was proposed recently that the capping box sequence can function as an effective helix stop signal in proteins (Harper & Rose, 1993) and in peptides (Dasgupta & Bell, 1993). It was also shown that, in peptides, the capping box sequence effectively suppresses N-terminal fraying and generally stabilizes the α-helix (Zhou et al., 1994). Nevertheless, it was also demonstrated that capping box sequences can adopt a helical conformation in the middle of a helical region if the residues have favorable side-chain−side-chain interactions for the stabilization of their α-helical conformation (Jimenez et al., 1994). The conclusion was drawn that the capping box is more likely to be a nucleation center for formation of an α-helix than a signal for termination of the helix.

However, there is one special case of a capping box sequence: Pro at N1. The presence of Pro at this position prevents the N-cap residue from adopting helix angles, while capping box angles are permitted. Additionally, there is the statistical observation that a strong preference exists for Asn at the N-cap position and Pro at the N1 position of α-helices in proteins (Richardson & Richardson, 1988). Therefore, we think that the motif (Ser/Thr/Cys/Asp/Glu/Asn/Gln)-Pro-Xaa-(Glu/Gln)-Xaa could play the role of a stop signal for α-helices in proteins. The preference for N-cap appears to depend on local interactions near N-termini and on long-range contacts within a protein structure.

Concluding Remarks. Our approach explains the helical propensities of analogs of NPY(11-36), NPY(12-36), and NPY(13-36), substituted at the N-cap position, as well as of other nonrelated peptides. Predicted values for the highest and the lowest helical propensities are in good agreement with experimental results and values for almost all middle-range propensities fall within the range of experimental error. The good correlation, at approximately the same level, between the experimental data and the results of calculations for several different sets of peptides indicates the applicability of the approach to peptides with a wide range of lengths and sequences.

Despite its relative simplicity, the analysis presented in this report revealed the nature of capping box stabilization and introduced three new factors: additional hydrogen bonding for Ser, Thr, and Cys; the influence of charged residues at N4; and N-capping interactions of the α-amino group with the side chain of the N4 residue. The values of the energy parameters for these interactions were mainly taken from experiments with peptides that were not related to the capping box. The use of these energies, in the statistical mechanics calculations without any refinement, yielded surprisingly high correlations with available experimental data.

The modifications of the original set of parameters presented in this study affect the AGADIR predictions for peptides with capping box sequences only. The prediction using the approach presented here showed much more accurate results for this type of peptide than when the original set of parameters was used. Additionally, these modifications change nothing in the AGADIR predictions for peptides without capping box sequences, which form the majority in the peptide database used for the refinement of the original AGADIR set of parameters. Therefore, the total correlation between the theoretical predictions made using the approach presented in this study and the experimental data for all peptides in the database must be better than that obtained with the original set of parameters.

One may note, however, that by introducing three new parameters into the theory the correlation with experimental data improves without being statistically significant. This

is true if an empirical set of parameters for a theory is refined to fit the experimental data. However, the factors discussed here are not empirical ones, and we did not refine their energy contributions. We just took into account the data obtained experimentally in other peptide and protein systems. It is unlikely that any theory that pretends to be valid could ignore factors that have been shown experimentally to play an important role.

ACKNOWLEDGMENT

The authors are grateful to Dr. L. Serrano for a critical reading of a preliminary draft of the manuscript and for the most recent version of the AGADIR computer program that was used for the calculations.

REFERENCES

Aqvist, J., Luecke, H., Quiocho, F. A., & Warshel, A. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 2026−2030.
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.
Cantor, C. R., & Schimmel, P. R. (1980) *Biophysical Chemistry. Part 1: The Conformation of Biological Macromolecules*, W. H. Freeman, New York.
Chakrabartty, A., Doig, A., & Baldwin, R. L. (1993a) *Proc. Natl. Acad. Sci. U.S.A. 90*, 11332−11336.
Chakrabartty, A., Kortemme T., Padmanabhan, S., & Baldwin, R. L. (1993b) *Biochemistry 32*, 5560−5565.
Dasgupta, S., & Bell, J. A. (1993) *Int. J. Pept. Protein Res. 41*, 499−511.
Doig, A., Chakrabartty, A., Klingler, T., & Baldwin, R. L. (1994) *Biochemistry 33*, 3396−3403.
el Masry, N. F., & Fersht, A. R. (1994) *Protein Eng. 7*, 777−782.
Finkelstein, A. V., Badretdinov, A. Y., & Ptisyn, O. B. (1991) *Proteins: Struct., Funct., Genet. 10*, 287−299.
Forood, B., Feliciano, E. J., & Nambiar, K. P. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 838−842.
Gans, P., Lyu, P. C., Manning, M., Woody, R., & Kallenbach, N. R. (1991) *Biopolymers 31*, 1605−1614.
Harper, E. T., & Rose, G. D. (1993) *Biochemistry 32*, 7605−7609.
Hol, W. G. J., van Duijnen, P. T., & Berendsen, H. J. C. (1978) *Nature 273*, 443−446.
Huyghues-Despointes, B. M. P., Scholtz, J. M., & Baldwin, R. L. (1993) *Protein Sci. 2*, 1604−1611.
Jimenez, M. A., Muñoz, V., Rico, M., & Serrano, L. (1994) *J. Mol. Biol. 242*, 487−496.
Lifson, S., & Roig, A. J. (1961) *J. Chem. Phys. 34*, 1963−1974.
Lockhart, D. J., & Kim, P. S. (1992) *Science 257*, 947−951.
Lockhart, D. J., & Kim, P. S. (1993) *Science 260*, 198−202.
Lyu, P. C.,Wemmer, D. E., Zhou, H. X., Pinker, R. J., & Kallenbach, N. R. (1993) *Biochemistry 32*, 421−425.
Muñoz, V., & Serrano, L. (1994) *Nature Struct. Biol. 1*, 399−409.
Muñoz, V., & Serrano, L. (1995a) *J. Mol. Biol. 245*, 275−296.
Muñoz, V., & Serrano, L. (1995b) *J. Mol. Biol. 245*, 297−308.
Muñoz, V., Blanco, F. J., & Serrano, L. (1995) *Nature Struct. Biol. 2*, 380−385.
Padmanabhan, S., & Baldwin, R. L. (1994) *J. Mol. Biol. 241*, 706−713.
Richardson, J. S., & Richardson, D. C. (1988) *Science 240*, 1648−1652.
Scholtz, J. M., & Baldwin, R. L. (1992) *Annu. Rev. Biophys. Biomol. Struct. 21*, 95−118.
Scholtz, J. M., Qian, H., York, E. J., Stewart, J. M., & Baldwin, R. L. (1991) *Biopolymers 31*, 1463−1470.
Seale, J. W., Srinivasan, R., & Rose, G. D., (1994) *Protein Sci. 3*, 1741−1745.
Serrano, L., & Fersht, A. R. (1989) *Nature 342*, 296−299.
Strickland, E. H. (1974) *CRC Crit. Rev. Biochem. 2*, 113−175.
Vasquez, M., & Scheraga, H. A. (1988) *Biopolymers 27*, 41−58.
Vasquez, M., Pincus, M., & Scheraga, H. A. (1987) *Biopolymers 26*, 351−371.
Yumoto, N., Murase, S., Hattori, T., Yamamoto, H., Tatsu, Y., & Yoshikawa, S. (1993) *Biochem. Biophys. Res. Commun. 196*, 1490−1495.
Zhou, H. X., Lyu, P.,Wemmer, D., & Kallenbach, N. R. (1994) *Proteins: Struct., Funct., Genet. 18*, 1−7.
Zimm, B. H., & Bragg, J. K. (1959) *J. Chem. Phys. 31*, 529−535.

BI9513766