Multiple Conformations and Proline *cis–trans* Isomerization in Salmon Calcitonin: A Combined Nuclear Magnetic Resonance, Distance Geometry, and Molecular Mechanics Study‡

Pietro Amodeo,§ Maria Antonietta Castiglione Morelli,‖ and Andrea Motta*,⊥

Dipartimento di Chimica, Università di Napoli Federico II, 80134 Napoli, Italy, Dipartimento di Chimica, Università della Basilicata, 85100 Potenza, Italy, and Istituto per la Chimica di Molecole di Interesse Biologico del CNR, 80072 Arco Felice (Napoli), Italy

Received April 18, 1994; Revised Manuscript Received June 15, 1994®

ABSTRACT: The relationship between multiple conformations and proline *cis–trans* isomerization in salmon calcitonin (sCT) has been investigated by ^{1}H-NMR, distance geometry, and molecular mechanics. In different solvents, the isomerization of Pro^{23} induces resonance heterogeneity for amino acids adjacent to it. NOESY experiments have related such heterogeneity to two different isoforms of the hormone, which interconvert into each other in polar solvents as well as in SDS micelles. The *cis–trans* ratio was found to depend upon the structure of the hormone. In water, the random-coil sCT showed a 35% *cis* population, while 16% *cis* was observed in the folded SDS-bound hormone. Such a decrease contributes, *per se*, a Gibbs free energy stabilization of 3.10 kJ mol^{-1}. Except for the 21–25 region, a common NOE pathway was found for both isomers. This is in agreement with the common secondary structure for both isoforms: a central helix and an extended C-terminal segment interacting with it. Calculations indicated that while the *trans* isomer is helical in the Thr^6–Tyr^{22} region, a shorter helix (Thr^6–Lys^{18}) is present in the *cis* isoform. It is concluded that isomerization of Pro^{23} does not alter the three-dimensional structure of the hormone, although *trans* structures show a lower average NOE root mean square deviation and a higher relative stability. Both *cis* and *trans* structures satisfactorily reproduce the experimental data, although they do not fulfill the whole set of NOE restraints, pointing out the danger involved in structure calculation whenever the *cis–trans* equilibrium does not affect the global fold.

A well-known facet of proline is its unique ability to form *cis* peptide bonds and undergo *cis–trans* isomerization. The equilibrium between *cis* and *trans* conformation of an Xaa–Pro peptide bond induces conformational heterogeneity and appears to be widespread in proteins. NMR[1] evidence has been gathered on the presence of multiple conformations in staphylococcal nuclease (Fox et al., 1986; Evans et al., 1987, 1989; Alexandrescu et al., 1988), insulin (Higgins et al., 1988), and calbindin D_{9k} (Chazin et al., 1989; Kördel et al., 1990). Conformational heterogeneity in native proteins has also been detected by X-ray analysis (Huber & Steigeman, 1974; Smith et al., 1986; Stewart et al., 1990; Svensson et al., 1992).

The presence of Pro-induced multiple conformations poses serious problems for the calculation of accurate three-dimensional solution structures from NOE data. If the conformations are all different, it is in theory possible to identify separate pathways of NOE connectivities for all of them and calculate the structure of each conformer. In isomers with structural perturbations localized around a Pro, most cross peaks of the different forms are degenerate and NMR restraints cannot be assigned specifically to each isoform. Furthermore, if two differently populated conformers are present in solution, the distance between two protons could be significantly shorter in the minor form than that in the major form, yet corresponding chemical shifts are the same. Thus a weak contribution from the major form cannot be distinguished from a strong contribution from the less abundant minor form. Clearly, then, the existence of a minor isoform linked to proline *cis–trans* isomerization in biomolecules must be taken into account in order to avoid fitting of NOE data in terms of a single family of conformations.

A way of circumventing this problem is to produce a mutant protein without the proline residue. However, from the structural point of view, a safe mutation would be to substitute proline with an amino acid which lacks the amide proton and forces the peptide bond into either the *cis* or the *trans* state only.

Here, we analyze the influence of proline *cis–trans* isomerism on the conformation of the hormone sCT. It is a 32 amino acid single-chain polypeptide hormone with an N-terminal disulfide bridge between residues 1 and 7 and a carboxyl-terminal prolinamide (Chart 1) (Breimer et al., 1988). sCT inhibits osteoclast activity and promotes Ca^{2+} uptake of bones (Azria, 1989). It is widely used as an anti-osteoporosis drug, but its mechanism of action still awaits a satisfactory explanation.

In this study our strategy is twofold. The first part reports NMR evidence that three-dimensional structure and local proline equilibrium are interdependent, although the structural perturbations accompanying the isomerization are very localized. The second part is an attempt to interpret the NOE data of SDS-bound sCT in terms of *cis* and/or *trans* conformers by using DG and EM calculations. The results (both the *cis* and the *trans* conformers satisfy the set of observed

‡ This paper is dedicated to the memory of Claudio Sellitti.
§ Università di Napoli Federico II.
‖ Università della Basilicata.
⊥ Istituto per la Chimica di Molecole di Interesse Biologico del CNR.
® Abstract published in *Advance ACS Abstracts,* July 15, 1994.

[1] Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; sCT, salmon calcitonin; SDS, sodium dodecyl sulfate; DG, distance geometry; EM, energy minimization; COSY, two-dimensional correlated spectroscopy; MLEV-17 TOCSY, total correlation spectroscopy obtained by using the MLEV-17 pulse sequence; NOESY, two-dimensional NOE spectroscopy; pH*, uncorrected glass-electrode reading; CT, calcitonin; $^3J_{HN\alpha}$, vicinal coupling constant between NH and α protons.

0006-2960/94/0433-10754$04.50/0

Chart 1: Amino Acid Sequence of sCT with Proline Residues in Boldface

Cys[1]-Ser-Asn-Leu-Ser[5]-Thr-Cys-Val-Leu-Gly[10]-Lys-Leu-
Ser-Gln-Gln[15]-Leu-His-Lys-Leu-Gln[20]-Thr-Tyr-**Pro**-
Arg-Thr[25]-Asn-Thr-Gly-Ser-Gly[30]-Thr-**Pro**-NH_2

NOEs) point out the dangers involved in the use of distances determined by NOE measurements without considering the presence of more than one conformation in solution.

MATERIALS AND METHODS

NMR Data Collection. sCT and its Leu[23]-Ala[24] analogue were prepared at Bachem Inc., Torrance, CA, by Dr. Nagana A. Goud. Sample concentration ranged from 0.5×10^{-3} to 5×10^{-3} M, the highest concentration being used for 95% ^1H_2O–5% ^2H_2O solution in presence of 0.90 M SDS. The detergent was purchased from Cambridge Isotope Laboratories (Woburn, MA). Isotopically labeled solvents [^2H_2O and $(C^2H_3)_2SO$] were obtained from Aldrich (Milwaukee, WI).

^{1}H-NMR spectra were recorded at 500 MHz on a Bruker AMX-500 spectrometer interfaced to an ASPECT X32 computer and referenced to sodium 3-(trimethylsilyl)[2,2,3,3-$^2H_4$]propionate or to the residual $(C^2H_3)_2SO$ signal. The relative concentrations of the *cis* and *trans* forms of the Tyr[22]–Pro peptide bond were obtained by integrating the calculated AA′XX′ aromatic spectrum of each conformer in all solvents. For this purpose we used PANIC, a Bruker LAOCOON-type program (Castellano & Bothner-By, 1964) running on an ASPECT-2000 computer. Two-dimensional experiments, namely, phase-sensitive double quantum filtered COSY (Piantini et al., 1982), MLEV-17 clean TOCSY (Braunschweiler & Ernst, 1983; Bax & Davis, 1985; Griesinger et al., 1988), chemical exchange, and NOESY experiments (Jeener et al., 1979; Macura & Ernst, 1980) were acquired by using the time-proportional phase incrementation scheme (Drobny et al., 1979). Usually, 512 equally spaced evolution time period t_1 values were acquired, averaging 16–64 transients of 2048 points, with 6024 Hz of spectral width. Time domain data matrices were zero filled to 4 K in both dimensions, thus yielding a digital resolution of 2.94 Hz/point. A Gaussian window in ω_2 and a 90°-shifted squared sine-bell window in ω_1 were applied before transformation. Chemical exchange and NOESY spectra were acquired with different mixing times. The *cis–trans* exchange rate in SDS was estimated at 325 K according to the method proposed by Jeener et al. (1979). For a system exchanging between two sites (*cis* and *trans*) the cross-peak amplitude goes through a maximum for

$$\tau_{mix} = \{\ln[(R_1 + k)/R_1]\}/k \quad (1)$$

τ_{mix} being the mixing time and k being the exchange-rate constant, and $R_1 = T_1^{-1}$. In order to achieve cross peak maximum intensity, exchange experiments were carried out with mixing times ranging between 0.50 and 1.2 s. A plot of cross-peak intensity *vs* mixing time presented a maximum for 0.85 s. Spin-lattice relaxation time was measured by selective inversion–recovery for the δ protons of Pro[23], obtaining a value of 0.95 s. According to the found experimental parameters, from eq 1 a value of 0.26 s^{-1} could be estimated for the rate of transfer at 325 K. Irradiation of the ^1H_2O resonance was achieved in the coherent mode (Zuiderweg et al., 1986).

Calculations. Calculations were performed as follows. (a) Two hundred sixty distance restraints (including the Thr[31]-NH–His[17]$CH_2\beta$, Asn[26]NH–Lys[18]$CH_2\delta$, and Asn[26]NH–Leu[19]-$CH\gamma$ long-range distances) were produced from NOESY spectra (mixing times, 0.060, 0.12, and 0.20 s) of sCT in SDS at 310 K, as previously described (Motta et al., 1991). They were used as upper limit distances to generate structures using the distance geometry program DIANA (Güntert et al., 1991). The variable target function was changed according to the standard strategy from level 1 to 32. The long-range restraints were included in the calculations either simultaneously or singularly, or they were not included at all, thus generating five sets of structures. In order to avoid that short distances in the calculated structures corresponded to nonobservable NOEs, some significant lower limits were imposed in calculations to reinforce the absence of observable NOEs. This gave rise to three additional sets of starting structures which included the long-range NOEs singularly. One hundred thirty structures were randomly generated for sCT with Pro[23] both in *cis* and in *trans* conformation. The same distance restraints were used as upper limits for both isoforms when they presented the same NOE, or spectral degeneracy prevented specific attribution. Only three NOEs involving residues 21–25 were different for the two isoforms. One hundred *cis* structures and 103 *trans* structures converged with small residual constraint violations and were chosen for further refinement.

(b) The converged structures were refined by EM, performed with the package AMBER, version 4.0 (Weiner & Kollman, 1981; Weiner et al., 1986). Since many NMR restraints cannot be assigned specifically to each isomer, minimization was achieved both without and with restraints (K_{dc} = 500 kJ nm^{-2} mol^{-1}). A conjugate gradient algorithm was used with a gradient convergence norm of less than 10^{-3} kJ mol^{-1} nm^{-1}. Except for the orientation of some side chains, no differences were found for the two sets of structures without and with NMR restraints. To account for the SDS anisotropic environment, we used in the minimizations two limiting values of the dielectric constant: $\epsilon = r$ (where r is the interatomic distance) for the hydrophobic bulk environment and $\epsilon = 10r$ for the polar environment (Polinsky et al., 1992). They gave rise to a pair of energy-minimized structures for each starting DG conformation. Although this approach does not predict the effects of specific sCT–micelle interactions, it allows a control on the stability of the calculated structures, eliminating those which are stable *in vacuo* only for a specific value of dielectric constant.

(c) Energy contributions from different molecular regions were evaluated, and relative stabilities for $\epsilon = r$ and $\epsilon = 10r$ were compared for both isoforms.

(d) NOE violations and ΔEs were used to select a set of "best" *cis* and *trans* structures, according to the following criteria: (1) the strict fulfillment of NOE restraints in the 1–8 region; (2) the loose fulfillment of the three long-range restraints; (3) (for the *trans* form only) the overall fitting of NOEs in the region 21–26, where the Pro[23] isomerism gives rise to the only univocally detectable spectral differences; and (4) the total energy differences in the $\epsilon = r$ calculations with a 320 kJ mol^{-1} threshold. Energetic criteria were used not to provide a scale of accurate relative stability but to look for possible correlations between stability and overall folding pattern. In fact, for both isoforms, no structure complying with the first three, for the *trans*, or the first two, for the *cis*, requirements exhibited an energy difference over the absolute minimum higher than the prefixed threshold. This identified two sets of best structures, 37 for the *cis* and 36 for the *trans* form.

(e) An accurate comparison of the selected *cis* and *trans* structures was performed in order to characterize the main conformational differences between the families.

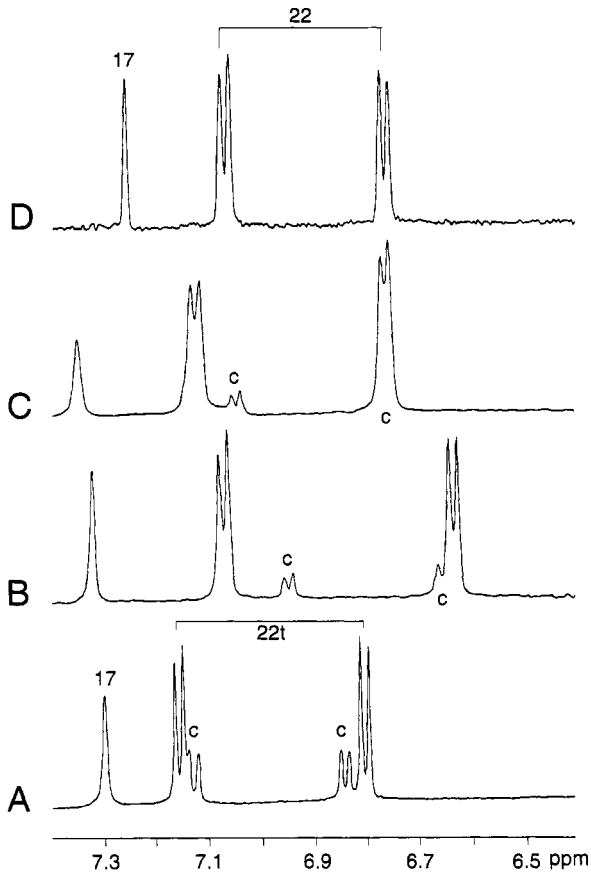

FIGURE 1: 500-MHz ^{1}H-NMR spectra of sCT dissolved in (A) ^2H_2O at 310 K and pH* 7.0, (B) 90% $(C^2H_3)_2SO$–10% ^2H_2O (v/v) at 278 K, and (C) aqueous SDS at 310 K and pH* 7.0, and of (D) Leu^{23}-Ala^{24}-sCT in ^2H_2O at 310 K and pH* 7.0. Residues are labeled by sequence numbers. t and c refer to *trans* and *cis* isomers, respectively.

RESULTS

Relationship between Three-Dimensional Structure and Local Proline Equilibrium. Well-resolved major and minor resonances were observed in the ^{1}H-NMR spectrum of sCT in several solvents. In Figure 1 are shown the spectra of the aromatic regions of sCT dissolved in ^2H_2O, in 90% $(C^2H_3)_2SO$–10% ^2H_2O (v/v), and in aqueous SDS, together with that of Leu^{23}-Ala^{24}-sCT (spectra A–D, respectively). The resonances of Tyr^{22} are accompanied by minor doublets (marked with a c) throughout spectra A–C. At 310 K we observed for them an intensity (with respect to the major peak) of 36% in ^2H_2O (spectrum 1A), 26% in $(C^2H_3)_2SO$ (not shown), and 15% in aqueous SDS (spectrum 1C), while 20% was observed in $(C^2H_3)_2SO$–^2H_2O at 278 K (spectrum 1B).

Regions of resonance heterogeneity were found around Pro^{23} after the sequential assignment was accomplished in all of the above solvent media by following the methodology outlined by Wüthrich (1986). The largest region of resonance heterogeneity (Thr^{21}–Thr^{25}) was observed in aqueous SDS.

The presence of a *cis* isomer for a Tyr^{22}–Pro^{23} bond was unequivocally ascertained by diagnostic NOESY cross peaks. The *cis* form allows much closer contacts between αCH(i) and αCH(i+1) and between NH(i) and αCH(i+1), whereas the *trans* form favors short distances between αCH(i) and δCH_2(i+1) and between NH(i) and δCH_2(i+1) (Wüthrich, 1986). In NOESY spectra (not shown) of sCT in $(C^2H_3)_2SO$–^2H_2O at 278 K and in SDS at 310 K (mixing times, 0.20 and 0.12 s, respectively), cross peaks between the αCH of Tyr^{22} and the δCH_2 of Pro^{23} and between the α protons of Tyr^{22} and Pro^{23} confirmed the presence of *trans* and *cis* isomers, respectively.

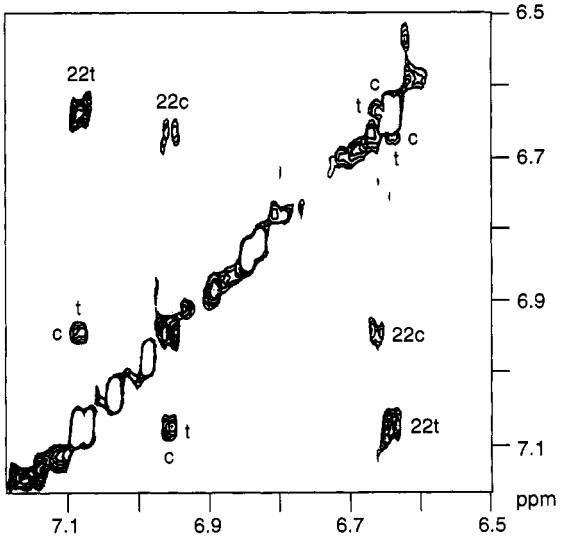

FIGURE 2: Exchange spectrum (τ_m = 0.65 s) of sCT in $(C^2H_3)_2SO$ at 310 K. Connectivities for the ring protons of Tyr^{22} *cis* and *trans* forms are indicated by 22c and 22t, respectively. Exchange cross peaks between the *cis* and *trans* forms are marked c and t, respectively.

The absence of resonance heterogeneity and of diagnostic NOEs for amino acids preceding Pro^{32} indicated that the amidated C-terminal proline is predominantly *trans*.

Figure 2 shows the low-field region of an exchange experiment of sCT in $(C^2H_3)_2SO$ at 310 K with a mixing time of 0.65 s. Together with the cross peaks between the ring protons of Tyr^{22} *trans* and *cis* (labeled 22t and 22c, respectively), exchange cross peaks are clearly observed at 7.08 and 6.96 and at 6.64 and 6.66 ppm between the *trans* (marked with a t) and *cis* (indicated with a c) forms. Exchange was also observed in SDS at 325 K, pH* 7.0 (data not shown). According to the method of Jeener et al. (1979), in aqueous SDS a value of 0.26 s^{-1} can be estimated for the rate of transfer from one form of the hormone to the other. Such a value is similar to the reported values for proline isomerism in proteins (Brandts et al., 1975).

This analysis provides conclusive evidence that the appearance of the interconverting forms of the hormone is the result of the *cis–trans* isomerization of Pro^{23}. The importance of Pro^{23} has also been suggested by Kern at al. (1993), who reported that in human CT the cytosolic peptidyl–prolyl *cis–trans* isomerase cyclophilin can accelerate catalytically the isomerization of the Pro^{23} peptide bond and not that of the Pro^{32} peptide bond.

The two species in solution were mapped by identifying separate pathways of sequential connectivities. As an example, Figure 3 reports the fingerprint region of a NOESY spectrum (mixing time, 0.10 s) of sCT in SDS showing sequential assignment for the polypeptide segment Gln^{20}–Asn^{26} both in the *trans* (solid lines) and in the *cis* (dashed lines) conformations. For the *trans* form, the αCH(i)–NH(i+1) NOE pattern (sequence numbers followed by a t) could easily be traced from Tyr^{22} down to Gln^{20}. The NH(i)–δCH_2(i+1) NOE between Tyr^{22} and Pro^{23}, typical for an Xaa–Pro *trans* peptide bond, and that between the δCH_2 of Pro^{23} and the NH of Arg^{24}, allowed assignment up to Asn^{26}. The *cis* isoform (sequence numbers followed by a c) shows a common NOE pathway with the *trans* conformation up to Gln^{20}. The alternative conformation starts at Thr^{21} as indicated by a weak αCH*trans*(i)–NH*cis*(i+1) cross peak linking Gln^{20} to Thr^{21}*cis*. A weak NH(i)–αCH(i+2) between Tyr^{22}*cis* and Arg^{24}*cis* allowed identification of the remaining *cis* pathway up to Thr^{25}*cis*. It was connected to Asn^{26} by an αCH*cis*(i)–NH*trans*(i+1) cross peak. An NOE between NH of Arg^{24}*cis*

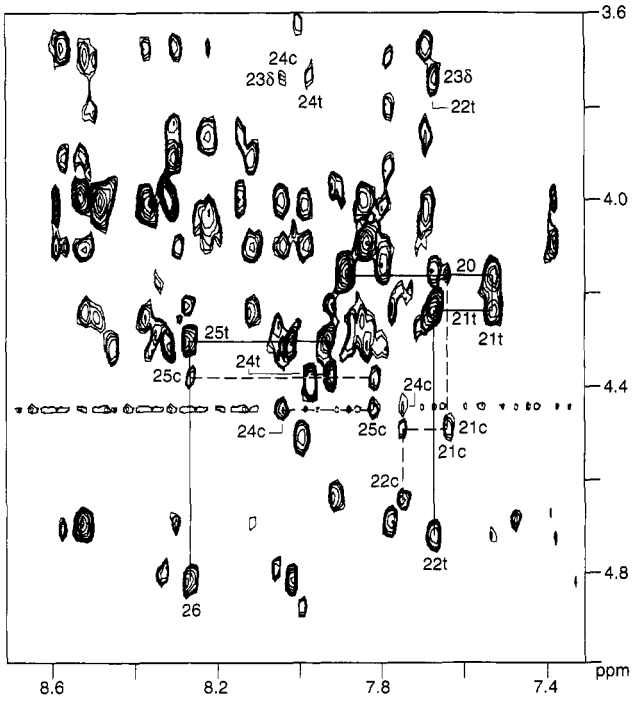

FIGURE 3: Fingerprint region of a NOESY spectrum (τ_m = 0.10 s) of sCT in 0.90 M SDS at 310 K and pH 3.7. The sequential assignment for the polypeptide segment Gln^{20}–Asn^{26} both in the *trans* (solid lines) and in the *cis* (dashed lines) conformation is shown.

and δCH_2 of Pro^{23} is also observed. A comparison of sequential NOEs shows no other differences for corresponding residues in the two isoforms.

$^3J_{HN\alpha}$ coupling constants for the *cis* isoform were estimated from a highly resolved double quantum filtered COSY experiment recorded in ^1H_2O. Except for Tyr^{22} and Arg^{24}, no significant differences ($\Delta^3J_{HN\alpha} > |2|$ Hz) were found for the other residues of the 21–25 segment when compared with the corresponding values of the *trans* isoform.

Stability of the Conformation with the trans Bond. In different solvents, sCT can assume structures of increasing complexity, in a sort of folding pathway. The random-coil structure in water becomes a short double-stranded antiparallel β sheet in a $(C^2H_3)_2SO$–water mixture at 278 K (Motta et al., 1989) because of the increased viscosity of the solvent (Motta et al., 1987; Amodeo et al., 1991). Such a structure evolves into an amphipathic α helix when in aqueous SDS (Motta et al., 1991). An α helix was also observed by NMR in trifluoroethanol (Meyer et al., 1991) and in methanol (Meadows et al., 1991). The increased complexity of the conformation is paralleled by a decrease of the *cis* population (*vide supra*). This reflects the stability of the conformation with the *trans* bond and implies that *cis–trans* isomerization is dependent on the global conformation of the hormone. An estimate of the influence of the three-dimensional structure on the local proline equilibrium can be obtained from the *cis* and *trans* populations in SDS (folded) and water (random coil) conformational states, although there are no physical transitions between these solvents.

Following Grathwohl and Wüthrich (1976a), we can calculate the free energy ΔG^0 for the *cis* = *trans* equilibrium of the Xaa–Pro peptide group:

$$\Delta G^0 = -RT \ln K; \quad K = [trans]/[cis] \qquad (2)$$

ΔG^0 is determined from the relative concentrations of the two species. It has been suggested (Wüthrich & Grathwohl, 1974) that

$$\Delta G^0 = \Delta G^0_{XP} + \Delta G^0_{conf} \qquad (3)$$

where ΔG^0_{XP} is a local term characteristic for the covalent structure of the fragment Xaa–Pro and ΔG^0_{conf} accounts for effects on *cis–trans* equilibrium arising from the occurrence of energetically preferred conformations.

The validity of the assumption that *cis* and *trans* populations in water can be used to evaluate ΔG^0_{XP} was verified at 295 K by comparing our value with those reported for model peptides. At 295 K, for sCT a 28% *cis* content was measured in water, from which we obtained a ΔG^0 value of −2.31 kJ mol^{-1}. Such a value correlates well with the ΔG^0_{FP} = −2.23 kJ mol^{-1} obtained for the protonated dipeptide H-Phe-Pro-OH, representative of a random-coil peptide in aqueous solution (Grathwohl & Wüthrich, 1976b).

By considering the relative intensities of the *trans* and *cis* peaks in water and SDS at 310 K, we can evaluate the change in the Gibbs free energy for the reduction of the *cis* population on going from water to SDS. From eq 2, we obtain −1.59 kJ mol^{-1} for water and −4.69 kJ mol^{-1} for SDS, and they can be considered representative of the random-coil and folded conformations of the hormone. From the above ΔG^0 values, we obtain a $\delta\Delta G^0$ of −3.10 kJ mol^{-1}, which is related to the stabilization afforded by the reduction of the *cis* population in the folded structure.

Calculation of Structures. In this section we investigate the possibility of reproducing the NOE pattern of SDS-bound sCT in terms of *cis* and/or *trans* isoforms.

In order to release steric strains, the well-converged 100 *cis* and 103 *trans* DIANA structures were subjected to EM. They had low target-function values and fulfilled the NOE distance restraints quite satisfactorily (maximum distance violation <0.05 nm). Since NMR restraints cannot be assigned specifically to each isoform, minimization was first achieved without restraints. Minimized structures without and with restraints were found to differ only for the orientation of some side chains. However, to ensure consistency with NOESY data, in the following calculations we explicitly used experimental restraints. For all of the DIANA structures, AMBER was able to find a low-energy conformation in the near vicinity, with only moderate increases of the sum of restraint violations (the average value of the sum of violations was 0.69 ± 0.10 nm before energy minimization and 1.0 ± 0.12 nm after energy minimization). Both *cis* and *trans* forms of sCT present a central α helix and an extended C-terminal segment folded back toward the helix. However, analysis of the ϕ and ψ angle distributions (not shown) indicates that most of the *trans* structures present values of ϕ and ψ consistent with an α helix in the region 6–22, while the *cis* form breaks the helix at Lys^{18}. For both sets of structures a larger ϕ and ψ scattering is observed in the N- and C-terminal regions.

No significant differences were found in the backbone of both isoforms for $\epsilon = r$ and $10r$. $\epsilon = r$ favors the formation of intramolecular salt bridges involving the side chains of Lys^{11}, Glu^{15}, and Lys^{18}; this finding only alters the order of the relative stability of the structures without modifying the set of structures.

The total energy of the two isoforms was evaluated by "factorizing" the molecule into three regions: head (residues 1–8), helix (residues 9–17), and tail (residues 18–32). We also assumed that the helix minimizes the interactions between head and tail. Figure 4A is a distribution plot of the total energy of both *cis* and *trans* forms for $\epsilon = r$. The height of the bars is proportional to the number of structures whose energy spreads in contiguous ranges 40 kJ mol^{-1} wide. The *trans* form is on average slightly more stable (−1672.34 kJ

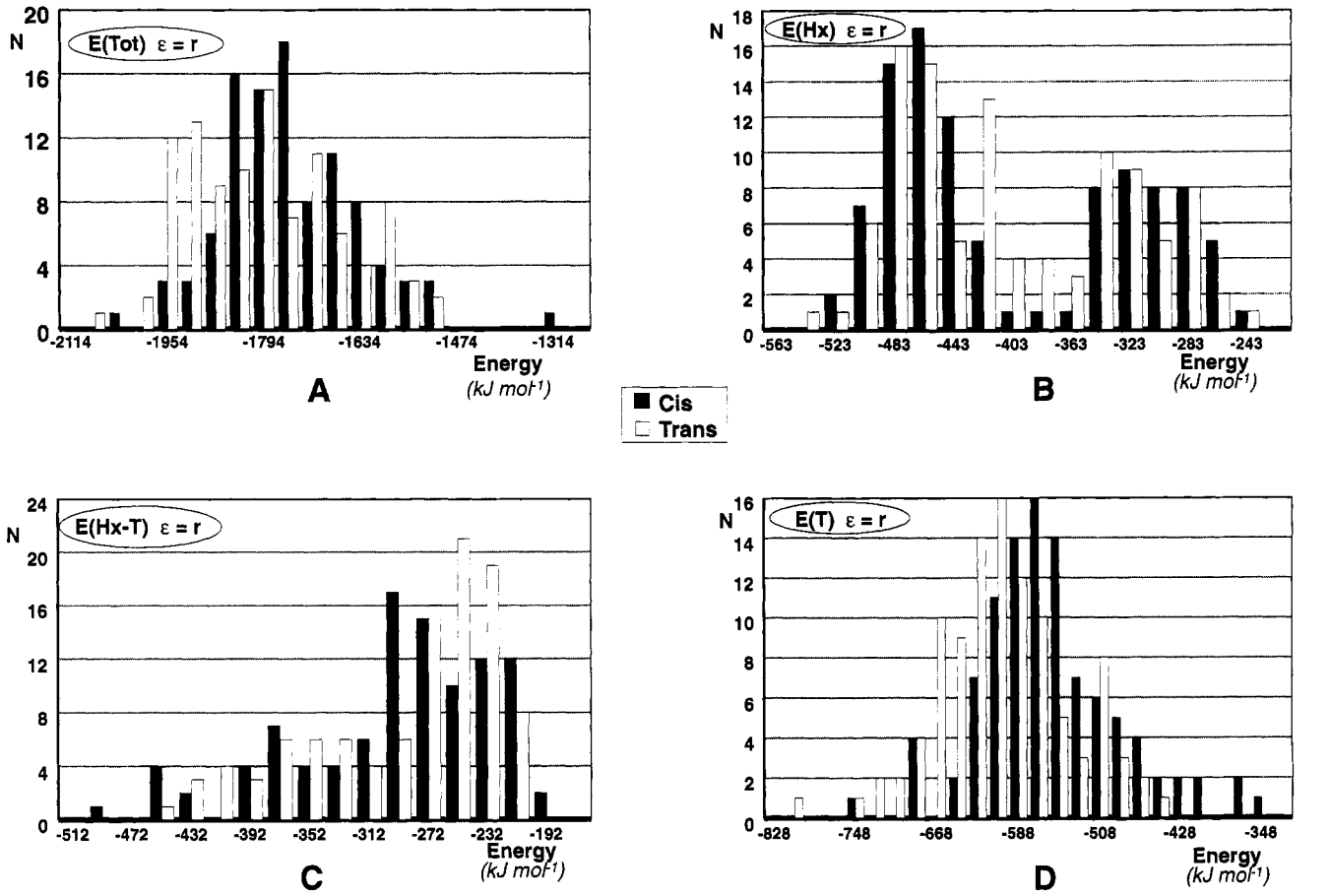

FIGURE 4: Energy distribution plots ($\epsilon = r$) for different subunits of the sCT calculated structures: (A) total energy, (B) energy of the helix region (residues 9–17), (C) helix–tail interaction energy, and (D) energy of the tail region (residues 18–32). N represents the number of structures with energies spanning a prefixed range (see text).

mol^{-1}) than the *cis* one (−1625.48 kJ mol^{-1}), and no contribution larger than 0.4 kJ mol^{-1} was calculated for the head-tail interaction. Only two *trans* structures (for $\epsilon = r$) showed a chain of hydrogen bonds and salt bridges joining the head and the tail, but their overall stability was very low [ΔE(tot) > 400 kJ mol^{-1} over the absolute minimum of the sets]. In all of the other cases the head and tail regions were completely independent, thus confirming the validity of the assumption.

Figure 4B shows the energy distribution of the helix calculated with a window of 20 kJ mol^{-1}. Both *cis* and *trans* forms present the same distribution, with two families of structures centered around average values of −460 and −290 kJ mol^{-1}, respectively. Such a large difference is ascribed to the formation of intrahelix salt bridges involving Lys^{11} and Glu^{15}. In the corresponding $\epsilon = 10r$ plot (not shown), the absence of salt bridges brings about the coalescence of the two families into a single bell-shaped distribution. Figure 4C reports the distribution of helix–tail interaction energies, calculated with a 20 kJ mol^{-1} window. While the average helix stability is similar for both *cis* and *trans* forms, *cis* structures exhibit more favorable helix–tail contributions. However, the tail energy distribution (Figure 4D) favors the *trans* structures. This analysis indicates that the *trans* form is on average more stable than the *cis* and that the higher stability of the *trans* form in the whole molecule may be mainly ascribed to the tail contribution.

In order to better characterize the stability of each isoform, we performed a thorough analysis of the backbone distance restraint violations. The analysis was carried out by using each DG restraint as an upper limit of 0.1-nm-wide range and defining a second range for distances less than ±0.05 nm outside the first range. The interproton distance values outside this second range were defined as "large" violations (indicated with a Ø in Figure 5), while those included in the second range were labeled as "small" (marked with a O in Figure 5).

Figure 5 shows the backbone distance restraint violations for selected *cis* and *trans* structures, as identified by the numbers in the leftmost column. Distance restraints were classified as intraresidue and interresidue, the latter referring to amino acids which are one, two, three, or four positions apart in the primary structure. Capital letters label restraints belonging to head (H), helix (X), and tail (T). In both isoforms, the helix essentially shows no violations for intra- and interresidue restraints, while few strong ones are found in the head of *trans* structures. Some systematic large violations are present in the tail region of the *cis* (Figure 5A) and *trans* (Figure 5B) structures. They can be explained with the high mobility of the C-terminal peptide in both isoforms. This results in a larger error deriving from the significant departure from the ideal isotropic molecular motion on which the restraint estimation is based.

Of the large violations observed in the tail region, highlighted by boldface capital letters in Figure 5, the *cis* structure exhibits a significantly higher number of violations. Figure 6 reports the percentage of violations of the most highly violated distance restraints for the best selected (see Materials and Methods) and the total number of structures of the two isoforms. Some of these violations (21–24, 23–26, 24–25, 24–26, and 25–26) concern NOEs which were considered to stem from the *trans* form. Although those NOEs are more frequently violated in the *cis* than in the *trans* form, an improved fitting of the 21–24, 23–26, and 24–25 NOEs is observed for selected *cis* structures. On the other hand, the violations of 21–24 and

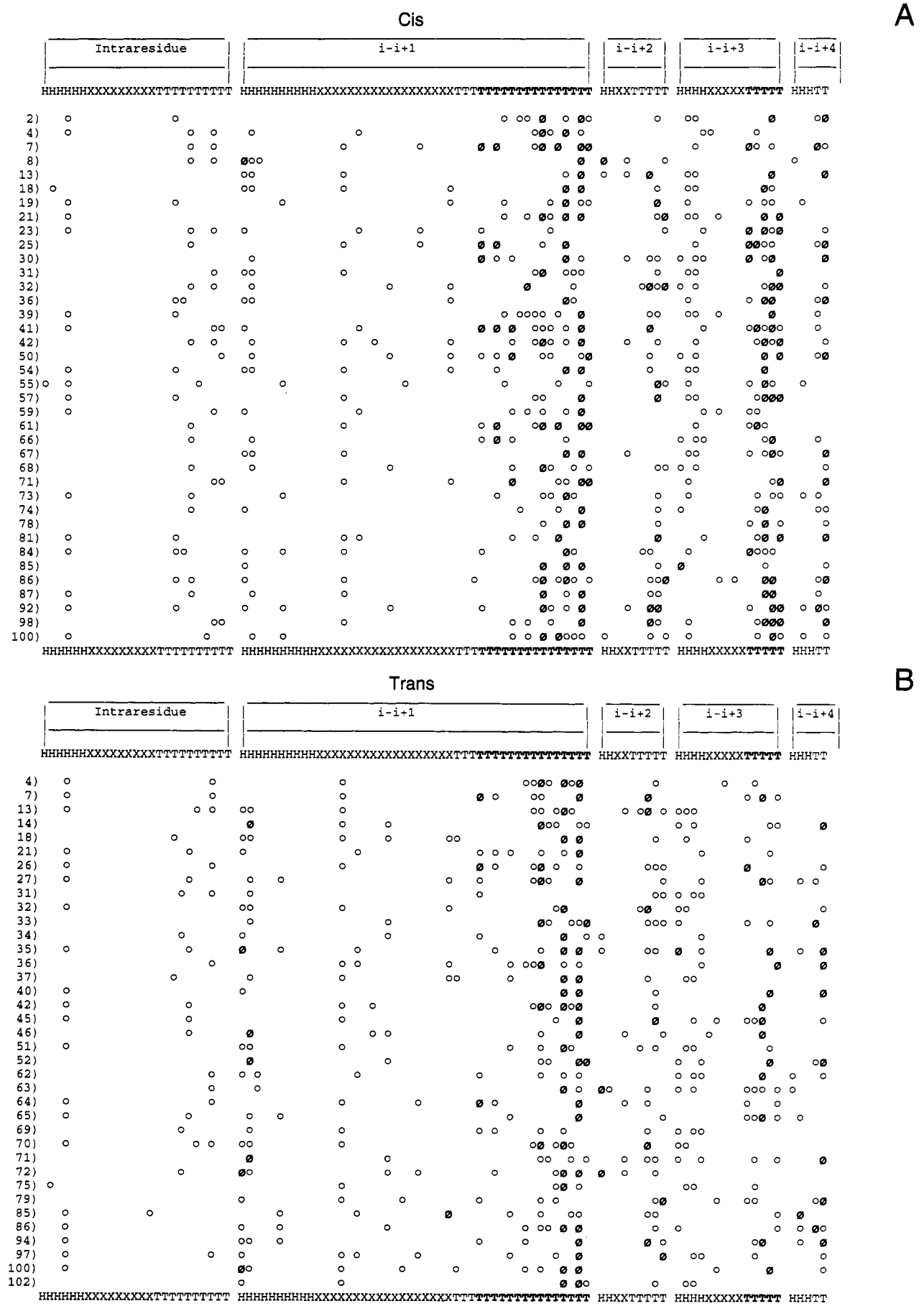

FIGURE 5: Symbolic plot of the backbone distance restraint violations for selected (A) *cis* and (B) *trans* SDS-bound sCT structures. Numbers in the left column refer to the position of each structure in the calculations. NOE restraints between protons located in residues separated by one to four positions along the sequence were classified as intra- or interresidue. Capital letters at the top and bottom label restraints belonging to head (H), helix (X), and tail (T). Boldface letters mark restraints more frequently violated in the *cis* form. Small violations are marked with a O, while large violations are marked with a Ø (for a definition of small and large violations, see text).

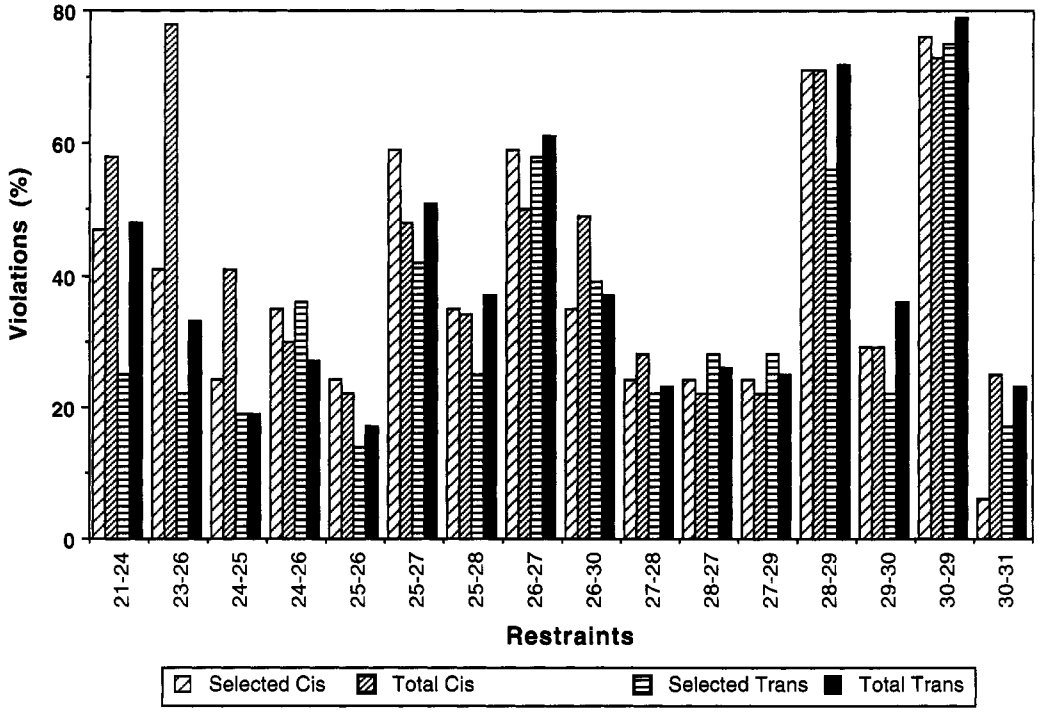

FIGURE 6: Plot of the percentage of violations of the most highly violated NOE distance restraints in SDS-bound sCT. Data are reported for the selected and total number of calculated *cis* and *trans* structures. Restraints are indicated with the sequence numbers of the two involved amino acids.

23–26 NOEs in the *trans* form are about half of the corresponding *cis* NOE violations even if selected sets are considered. Accordingly, it cannot be inferred that the above NOEs exclusively belong to the *trans* form.

An opposite trend is observed for the 24–26, 25–26, 25–27, and 25–28 restraints of the *cis* form, for which the selected structures show an increased fraction of violations. On going from total to selected sets, *trans* structures show smaller percentage variations than *cis* ones. The violation decrease detected for 21–24 and 23–26 NOEs in the selected *trans* form cannot be interpreted as for the *cis* form, since they have been used as selection criteria (see Materials and Methods). It appears that in the 21–27 region only the 24–26 restraint exhibits a significant difference in the violation percentage, although it is less than 10%, and the percentage of violations in the *trans* selected structures is less than 40%.

The C-terminal region involving residues 28–32 shows two large violations for both isoforms, namely, 28–29 and 30–29 NOEs with more than 70% of violations. No significant improvement is observed when selected *trans* structures are considered. This is justified by the fact that the tail is the most flexible region of the molecule, and its mobility accounts for the observed long-range NOEs. In addition, the mobility implies that the sampled sets of structures for the tail correspond to a significantly smaller fraction of available conformational space than in the rest of the molecule.

DISCUSSION

Minor population of the Tyr^{22}-Pro^{23} *cis* peptide bond was observed in different conformational states of sCT and was studied by using the well-resolved Tyr ring resonances as a specific probe. NMR evidence indicates that in sCT structural changes imposed by *cis–trans* isomerization of Pro^{23} are localized to the region around this residue and that the *cis–trans* equilibrium population is regulated by conformational transitions of the hormone. In fact, the *cis* isomer was found to be 36% for the random-coil hormone in water and 15% in the totally folded state in SDS. We estimated for the helical conformation with a *trans* bond a 3.10 kJ mol^{-1} stabilization relative to the decrease of the *cis* population from 36% to 15%. Calculations interpreted the greater stability of the *trans* form of SDS-bound sCT as originating from a longer average helical segment covering the 6–22 region. The values of ϕ and ψ of *cis* structures were consistent with a helix in the 6–18 region.

The dependence of the *cis–trans* population upon the tertiary structure has also been observed in larger protein systems. In ribonuclease A, the Tyr^{92}–Pro^{93} bond is *ca.* 70% *cis* in the unfolded protein and nearly 100% *cis* in the native state (Lin & Brandts, 1983). The restoration of enzymatic activity of the protein during refolding completely parallels the restoration of the *cis* form for Pro^{93} (Lin & Brandts, 1983). In this case the protein requires a *cis* bond in its biologically active conformation, and thus the refolding brings about a reduction of the *trans* population.

It has been suggested that *cis–trans* isomerism may contribute significantly to the function of many transmembrane transport proteins (Brandl & Deber, 1986). Such a process could alter the alignment of the transmembrane segment and thereby provide the conformational basis for a dynamic reversible mechanism for channel "opening–closing". This hypothesis rests on the fact that *trans* and *cis* peptide bond containing structures should be able to provide, in principle, the molecular basis for local conformational transitions about Pro sites in an environment of low dielectric constant. Exchange experiments indicated that *cis–trans* isomerization in sCT also takes place in a 0.90 M solution of SDS, used to mimic the membrane environment. Thus molecules as small as sCT do have the inherent propensity to exist in multiple conformational states of competitive energy which can interconvert into each other. However, our calculations showed that Pro^{23} *cis–trans* isomerism does not reverse the alignment of the backbone, but it shortens the length of the preceding helix and preserves the overall molecular shape of sCT.

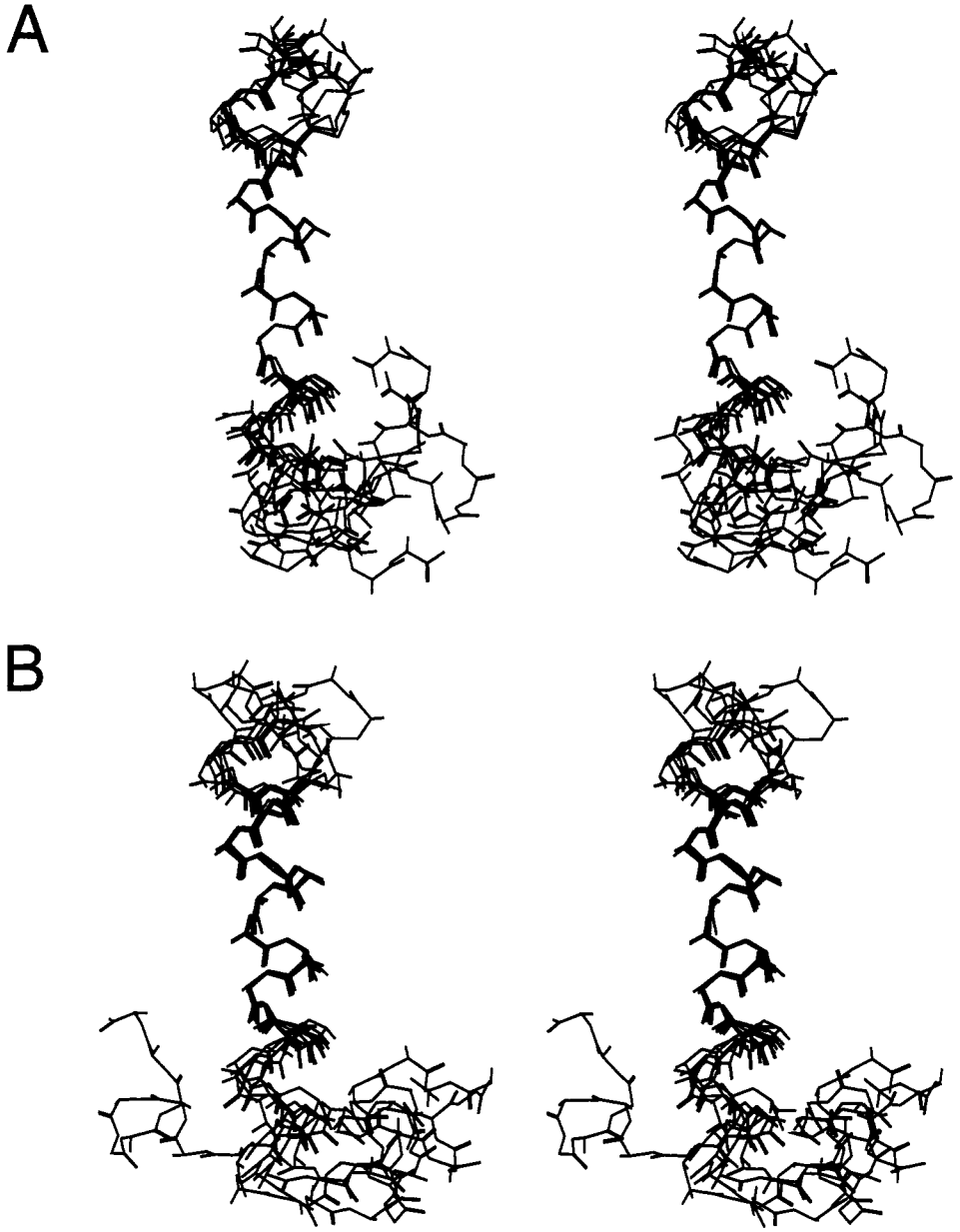

FIGURE 7: Stereoviews of five representative *cis* (A) and *trans* (B) sCT structures. They were superimposed for pairwise minimum root mean square differences of the backbone atoms of residues 10–16 with the lowest energy structure. For clarity, only backbone atoms are shown and the same orientation of the central helical region is used for the two sets.

The experimental NOE pattern is satisfactorily reproduced, although neither the *trans* nor the *cis* structure fulfills the whole set of NOE distance restraints. The main differences between *cis* and *trans* forms are the shorter central helix and the overall arrangement of the long-range NOEs. The family of *trans* structures showed a lower average NOE root mean square deviation and a higher relative stability. The *trans* family accounts for the observed NOEs around Pro^{23} and only two of the long-range ones, namely, $Asn^{26}NH$–Lys^{18}-$CH_2\delta$ and $His^{17}CH_2\beta$–$Thr^{31}NH$. These results are in agreement with those obtained by means of restrained molecular dynamics simulations with both "static" and time-averaged NMR distance restraints (Castiglione Morelli et al., 1992). On the other hand, several *cis* structures, although of a lower relative stability, were able to simultaneously fulfill two or (more loosely) three long-range restraints. Therefore, on the basis of the currently available experimental data, neither conformational freedom nor the long-range NOEs, induced by the tail folding, are diagnostic criteria to differentiate *cis* and *trans* isoforms.

Figure 7 shows a comparison between selected *cis* and *trans* structures. Due to the shorter helix and the consequent fulfillment of all of the long-range contacts, the *cis* isoform (Figure 7A) is more compact than the *trans* one (Figure 7B), although the overall molecular shape is preserved in both conformers. It is tempting to speculate that the *cis* isoform, more than the *trans* one, is involved in the interaction with the receptor, through the shorter helix. This is suggested by the observation in SDS of the residual 15% *cis* population and by the existence of interconverting, multiple conformational states of competitive energy. Accordingly, the isomerization would be required to reduce the length of the helix and to preserve the molecular flexibility during the interaction with the receptor.

ACKNOWLEDGMENT

We would like to thank Dr. Nagana A. Goud (Bachem Inc., Torrance, CA) for the samples of calcitonin used in this study.

REFERENCES

Alexandrescu, A. T., Mills, D. A., Ulrich, E. L., Chinami, M., & Markley, J. L. (1988) *Biochemistry 27*, 2158–2165.

Amodeo, P., Motta, A., Picone, D., Saviano, G., Tancredi, T., & Temussi, P. A. (1991) *J. Magn. Reson. 95*, 201–207.

Azria, M. (1989) *The Calcitonins: Physiology and Pharmacology*, Karger, Basel, Switzerland.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–366.

Brandl, C. J., & Deber, C. M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 917–921.

Brandts, J. F., Halvorson, H. R., & Brennan, M. (1975) *Biochemistry 14*, 4953–4963.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.

Breimer, L. H., MacIntyre, I., & Zaidi, M. (1988) *Biochem. J. 255*, 377–390.

Castellano, S., & Bothner-By, A. A. (1964) *J. Chem. Phys. 41*, 3863–3871.

Castiglione Morelli, M. A., Pastore, A., & Motta, A. (1992) *J. Biomol. NMR 2*, 335–348.

Chazin, W. J., Kördel, J., Drakenberg, T., Thulin, E., Brodin, P., Grundström, T., & Forsén, S. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 2195–2198.

Drobny, G., Pines, A., Sinton, S., Weitkamp, D., & Wemmer, D. (1979) *Faraday Symp. Chem. Soc.* No. 13, 49–55.

Evans, P. A., Dobson, C. M., Hatfull, G., Kautz, R. A., & Fox, R. O. (1987) *Nature (London) 329*, 266–268.

Evans, P. A., Kautz, R. A., Fox, R. O., & Dobson, C. M. (1989) *Biochemistry 28*, 362–370.

Fox, R. O., Evans, P. A., & Dobson, C. M. (1986) *Nature (London) 320*, 192–194.

Grathwohl, Ch., & Wüthrich, K. (1976a) *Biopolymers 15*, 2025–2041.

Grathwohl, Ch., & Wüthrich, K. (1976b) *Biopolymers 15*, 2043–2057.

Griesinger, C., Otting, G., Wüthrich, K., & Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870–7872.

Güntert, P., Braun, W., & Wüthrich, K. (1991) *J. Mol. Biol. 217*, 517–530.

Higgins, K. A., Craik, D. J., Hall, J. G., & Andrews, P. R. (1988) *Drug Des. Delivery 3*, 159–170.

Huber, R., & Steigeman, W. (1974) *FEBS Lett. 48*, 235–237.

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.

Kern, D., Drekenberg, T., Wikström, M., Forsén, S., Bang, H., & Fischer, G. (1993) *FEBS Lett. 323*, 198–202.

Kördel, J., Forsén, S., Drakenberg, T., & Chazin, W. J. (1990) *Biochemistry 29*, 4400–4409.

Lin, L.-N., & Brandts, J. F. (1983) *Biochemistry 22*, 559–563.

Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95–117.

Meadows, R. P., Nikonowicz, E. P., Jones, C. R., Bastian, J. W., & Gorenstein, D. G. (1991) *Biochemistry 30*, 1247–1254.

Meyer, J.-P., Pelton, J. T., Hoflack, J., & Saudek, V. (1991) *Biopolymers 31*, 233–241.

Motta, A., Picone, D., Tancredi, T., & Temussi, P. A. (1987) *J. Magn. Reson. 75*, 364–370.

Motta, A., Castiglione Morelli, M. A., Goud, N. A., & Temussi, P. A. (1989) *Biochemistry 28*, 7998–8002.

Motta, A., Pastore, A., Goud, N. A., & Castiglione Morelli, M. A. (1991) *Biochemistry 30*, 10444–10450.

Piantini, U., Sørensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800–6801.

Polinsky, A., Goodman, M., Williams, K. A., & Deber, C. M. (1992) *Biopolymers 32*, 399–406.

Smith, J. L., Hendrickson, W. A., Honzatko, R. B., & Sheriff, S. (1986) *Biochemistry 25*, 5018–5027.

Stewart, D. E., Sarkar, A., & Wampler, J. E. (1990) *J. Mol. Biol. 214*, 253–260.

Svensson, L. A., Thulin, E., & Forsén, S. (1992) *J. Mol. Biol. 223*, 601–603.

Weiner, P. K., & Kollman, P. A. (1981) *J. Comput. Chem. 2*, 287–303.

Weiner, P. K., Kollman, P. A., Nguyen, D. T., & Case, D. A. (1986) *J. Comp. Chem. 7*, 230–252.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.

Wüthrich, K., & Grathwohl, Ch. (1974) *FEBS Lett. 43*, 337–340.

Zuiderweg, E. R. P., Hallenga, K., & Olejniczak, E. T. (1986) *J. Magn. Reson. 70*, 336–343.