Journal of Biomolecular NMR, **13:** 161–174, 1999.
KLUWER/ESCOM

Conformational flexibility in calcitonin: The dynamic properties of human and salmon calcitonin in solution

Pietro Amodeo[a], Andrea Motta[a,*], Giuseppe Strazzullo[a] & Maria A. Castiglione Morelli[b]
[a]*Istituto per la Chimica di Molecole di Interesse Biologico (Istituto Nazionale di Chimica dei sistemi Biologici) del CNR, I-80072 Arco Felice (Napoli), Italy;*
[b]*Dipartimento di Chimica, Università della Basilicata, I-85100 Potenza, Italy*

Received 2 July 1998; Accepted 4 November 1998

Key words: calcitonin, conformational flexibility, DYANA, helix chirality, molecular dynamics, torsion-angle dynamics

Abstract

We have studied the dynamic properties of human (h) and salmon (s) calcitonin (CT) in solution. For both hormones, distance geometry in torsion-angle space has been used to generate three-dimensional structures consistent with NMR data obtained in sodium dodecyl sulfate micelles. For sCT and hCT we used, respectively, 356 and 275 interproton distances together with hydrogen-bonds as restraints. To better characterize their flexibility and dynamic properties two fully unrestrained 1100-ps molecular dynamics (MD) simulations in methanol were performed on the lowest-energy structures of both hormones. Statistical analyses of average geometric parameters and of their fluctuations performed in the last 1000 ps of the MD run show typical helical values for residues 9–19 of sCT during the whole trajectory. For hCT a shorter helix was observed involving residues 13–21, with a constant helical region in the range 13–19. Angular order parameters S(ϕ) and S(ψ) indicate that hCT exhibits a higher flexibility, distributed along the whole chain, including the helix, while the only flexible amino acid residues in sCT connect three well-defined domains. Finally, our study shows that simulated annealing in torsion-angle space can efficiently be extended to NMR-based three-dimensional structure calculations of helical polypeptides. Furthermore, provided that a sufficient number of NMR restraints describes the system, the method allows the detection of equilibria in solution. This identification occurs through the generation of 'spurious' high-energy structures, which, for right-handed α-helices, are likely to be represented by left-handed α-helices.

Introduction

Calcitonin (CT) is a single-chain polypeptide hormone of 32 amino acids with an N-terminal disulfide bridge between positions 1 and 7 and a C-terminal proline amide residue. Secreted by the parafollicular or C cells located in the thyroid glands of mammals and by the ultimobranchial gland in lower animals, CT inhibits osteoclastic bone resorption, induces calcium uptake from body fluids and plays an important role in calcium-phosphorus metabolism (Austin and Heat, 1981). In general, fish calcitonins have been found to be more potent biologically than mammalian ones. Three native calcitonins, salmon (sCT), porcine CT and human (hCT), and an analogue of eel CT are used therapeutically mainly for the treatment of osteoporosis, Paget's disease, and hypercalcemia of malignancy (Azria, 1989). However, sCT, differing from hCT in 16 of the 32 amino acids, is reported to produce antibodies in a significant number of patients (Levy et al., 1988), while hCT easily associate and precipitate as insoluble fibrils upon storing in aqueous solution, which is a defect in therapeutic use. In order to prepare new high potency molecules it would be desirable to find correlation between biological functions and structural features, but the molecular aspects of CT activity are not completely understood. A comparison of the sequences of several calcitonins (Dayhoff,

*To whom correspondence should be addressed.

1978) shows that only nine residues are common to all species so far studied (residues 1, 3–7, 9, 28 and 32) and they are essentially clustered in the N-terminal segment (Scheme 1). This suggests that differences in pharmacological potency have to be related to structural differences in the central and C-terminal regions. In particular, most of the amino acid substitutions (10 out of 16, in bold in Scheme 1) are concentrated in the segment 8–22, which is expected to form the amphipathic helix. NMR and distance geometry (DG) studies in a trifluoroethanol-water mixture (Meyer et al., 1991) and in methanol (Meadows et al., 1991) indicate that sCT indeed assumes an amphipathic α-helix in the region 8–22, while in sodium dodecyl sulfate (SDS) (Motta et al., 1991b; Castiglione Morelli et al., 1992) the helix spans the segment 6–22. For hCT, the 9–22 residue region has been reported to adopt a helical conformation in a trifluoroethanol-water mixture (Doi et al., 1990), while a shorter helix spanning the region 9–16 has been observed in SDS (Motta et al., 1998). Instead, in the so-called cryomixtures (Motta et al., 1987), both sCT (Motta et al., 1989) and hCT (Motta et al., 1991a) assumed an extended conformation with a short double-stranded antiparallel β-sheet in the central regions.

The presence of an amphipathic α-helix in the central region of CT has been reported to be important for the interaction with lipids and for hypocalcemic activity (Epand et al., 1983). Kaiser and co-workers (Moe and Kaiser, 1985; Green III et al., 1987) provided evidence that the model of an amphipathic α-helical region in residues 8–22 is a useful guide to designing potent sCT activity, although other studies suggest that conformational flexibility is also an important parameter for activity (Epand et al., 1986).

In order to evaluate the intrinsic molecular flexibility we have studied the solution structures of hCT and sCT in SDS, an environment employed to simulate the CT conformation in the presence of membrane-bound receptors. DG in torsion-angle space has been used to generate three-dimensional structures of both hormones consistent with NMR data. The 13 highest energy DG structures of hCT present an inversion of the helix chirality, although the amino acids were all found in the L configuration. On the contrary, the correct chirality of the α-helix was preserved in sCT. The use of the distance-geometry algorithm DIANA, which does not use torsion-angle dynamics, did not show helix chirality inversion for both hormones. Considering the number and the distribution of restraints along the helical chain and the efficiency of the torsion-angle molecular dynamics (Güntert and Mumenthaler, 1997; Stein et al., 1997), the presence of the inverted chirality in hCT seems to depend upon the looseness of the nuclear Overhauser enhancement (NOE) restraints, which originates from the intrinsic flexibility of the human form over the salmon one. Conformational differences in both hormones were evaluated by using unrestrained molecular dynamics (MD) simulations in methanol, which is known to induce both in human and salmon calcitonins secondary structures very similar to those found in detergent micelles (Epand et al., 1988; Meadows et al., 1991; Arvinte and Drake, 1993), but avoids the complexity of the micellar system in the calculations. Compared with sCT, the human form is found to have a shorter helix length and no helix–tail interactions. Analysis of the angular order parameter (S) of ϕ and ψ dihedral angles indicates the presence of three domains in the sCT structure with flexible loops connecting the head and the helix, and the helix and the tail regions, while for hCT the flexibility is extended over the whole polypeptide chain.

Finally, our results point to the conclusion that MD in torsion-angle space is a highly efficient method for structure calculations not only for proteins and nucleic acids (Güntert and Mumenthaler, 1997; Stein et al., 1997), but also for helical polypeptides. We have also found that, due to the accuracy of the sampling, it also appears to be able to uncover conformational equilibria in solution by detecting spurious high-energy structures.

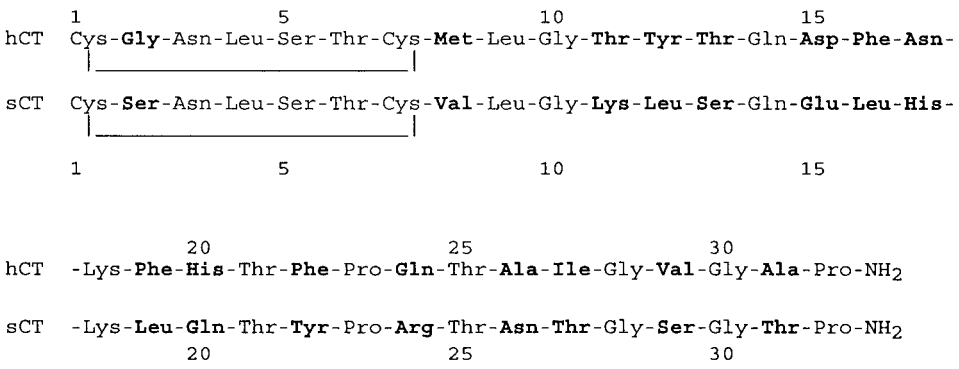
Materials and methods

Peptide synthesis

sCT and hCT were prepared by classical methods of peptide synthesis in solution and belonged to batches already described (Amodeo et al., 1994; Motta et al., 1998).

NMR data collection

For acquisition of NMR spectra, 5×10^{-3} M solutions of sCT and hCT in 95% ^1H_2O/5% ^2H_2O (v/v), pH 4.1 and 7.2 or in 100% ^2H_2O, pH* 3.7 and 6.8, all in phosphate buffer at 310 K, were used. Isotopically labeled solvents originated from Aldrich (Milwaukee, WI). A solution of perdeuterated SDS was added to a solution of a hormone, and the final mixture remained transparent. The concentration of SDS was maintained



Scheme 1. Amino acid sequence of hCT and sCT. Differences in primary structures are in boldface.

well above the critical micelle concentration, with a final CT-SDS ratio of ca. 1:120. Perdeuterated SDS was obtained from Cambridge Isotope Laboratories (Woburn, MA).

In order to obtain comparable data sets for both hormones, NMR experiments on sCT, although previously reported (Motta et al., 1991b), were again acquired in parallel with those on hCT. ^{1}H NMR spectra of both hormones were recorded on a Bruker DRX-500 spectrometer operating at 500 MHz using an inverse multinuclear probehead fitted with gradients along the X-, Y-, and Z-axes, and equipped with an SGI computer. Spectra were referenced to sodium 3-(trimethylsilyl)-[2,2,3,3-$^2H_4$]propionate. Homonuclear DQF-COSY (Piantini et al., 1982), clean TOCSY (Braunschweiler and Ernst, 1983; Griesinger et al., 1988) and NOESY (Jeener et al., 1979) spectra were recorded by using time-proportional phase incrementation of the first pulse (Drobny et al., 1979), and incorporating the WATERGATE sequence (Piotto et al., 1992) for water suppression. The gradient pulses were each 1 ms in duration and shaped to a sine envelope. 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 all zero-filled to 4K in both dimensions, thus yielding a digital resolution of 2.94 Hz/pt. Prior to Fourier transformation, a Lorentzian-Gaussian window with different parameters was applied for both t_1 and t_2 dimensions for all the experiments. NOESY spectra were obtained with different mixing times (0.10, 0.15 and 0.20 s), TOCSY experiments were recorded with a spin-lock period of 0.064 and 0.096 s, achieved with the MLEV-17 pulse sequence (Bax and Davis, 1985).

Slowly exchanging protons at pH* 7.0 were identified by recording a NOESY spectrum (0.20 s mixing time) of the peptides immediately after dissolution in ^2H_2O. For isolated resonances, measurements of $^3J_{HN\alpha}$ coupling constants were obtained from one-dimensional experiments acquired with 128 K points and application of strong Lorentzian-Gaussian resolution enhancement. For overlapping lines coupling constants were estimated from DQF-COSY spectra so that they are considered as apparent values (Neuhaus et al., 1985).

Processing was performed using TRIAD (Tripos Inc., St. Louis, MO) and AURELIA 2.0 (Neidig et al., 1995) (Bruker) running on Silicon Graphics computers.

Distance restraint determination

An initial list of distances was obtained for both calcitonins from a 0.10 s mixing time NOESY spectrum recorded at 310 K. Long-range NOEs, if present, were obtained from a 0.20 s NOESY experiment. NOEs volumes were integrated and calibrated with the AURELIA software (Neidig et al., 1995). For sCT we collected manually 356 NOE-derived restraints (154 intraresidual, 199 sequential and medium-range, and 3 long-range) and we used the intraresidual distance αCH_i-NH_i of Leu^9 for the calibration. This is an expanded and more accurate data set of restraints with respect to the semi-quantitative one previously used (Motta et al., 1991b; Castiglione Morelli et al., 1992). For hCT we collected 275 unambiguous distance restraints (143 intraresidual, and 132 sequential and medium-range). Analysis of the NOE distance restraints using the program DIANA (Güntert et al., 1991) revealed that some intraresidual interproton distance restraints (namely, αCH_i-NH_i, βCH_i-NH_i, αCH_i-βCH_i) placed no restriction on the structure as the upper distance limits imposed were smaller than

those allowed by the covalent geometry. Then we decided to differently calibrate intraresidual and interresidual NOEs. Interresidual NOEs were calibrated using as reference a sequential NH_i-NH_{i+1} distance corresponding to residues Tyr^{12}-Thr^{13} for which unambiguous experimental evidence of helical conformation exists. For each kind of intraresidual NOE we identified the allowed distance range from a systematic dihedral search using the AMBER standard covalent geometry (Weiner and Kollman, 1981). Distances were calibrated by using the Leu^9 αCH_i-NH_i NOE in the helical region, and then scaled to fit the corresponding allowed range. A $\pm$ 0.01 nm tolerance was allowed for each restraint. The β-methylene groups were stereospecifically assigned with the program HABAS (Güntert et al., 1989). A total of 12 β-methylene groups out of 23 for sCT (5, 9, 11, 12, 13, 14, 15, 17, 18, 19, 20 and 22), and 7 out of 20 β-methylene groups for hCT (residues 7, 8, 9, 12, 14, 19 and 24) were stereospecifically assigned. When no stereospecific assignment was possible for methyl and methylene protons, distance restraints were corrected for pseudoatom representation (Wüthrich et al., 1983). Hydrogen bonds were imposed between pairs of residues for which standard intrahelical αCH_i-NH_{i+3} and αCH_i-NH_{i+4} NOEs were observed for a slow exchanging amide proton (Wagner et al., 1987). For each hydrogen bond, two upper limit restraints were used between the NH-O (0.25 nm) and the N-O (0.35 nm) atom pairs. The Cys^1-Cys^7 disulfide bond was fixed by imposing a range of 0.20–0.21 nm on the S-S distance and a range of 0.30–0.35 nm on the S_γ-C_β distances.

Structure calculations

Structures were generated by the DG program DYANA (Güntert et al., 1997). A total of 390 structures of hCT, in 13 sets of 30 conformers each, were generated using different combinations of the MD time steps (4000, 6000, 8000, 10 000 and 20 000) and high temperatures of the annealing ($T_{high} = 6, 8, 10$). Thirty sCT structures were calculated by simulated annealing starting with a total of 8000 MD steps and $T_{high} = 6$. All of the other default values of temperature and time steps were used. Thirty hCT and sCT structures were also generated with the program DIANA (Güntert et al., 1991), using the standard strategy for the minimization of the target function from level 1 to 32. Structures of each CT were further refined by restrained energy minimization (EM) performed

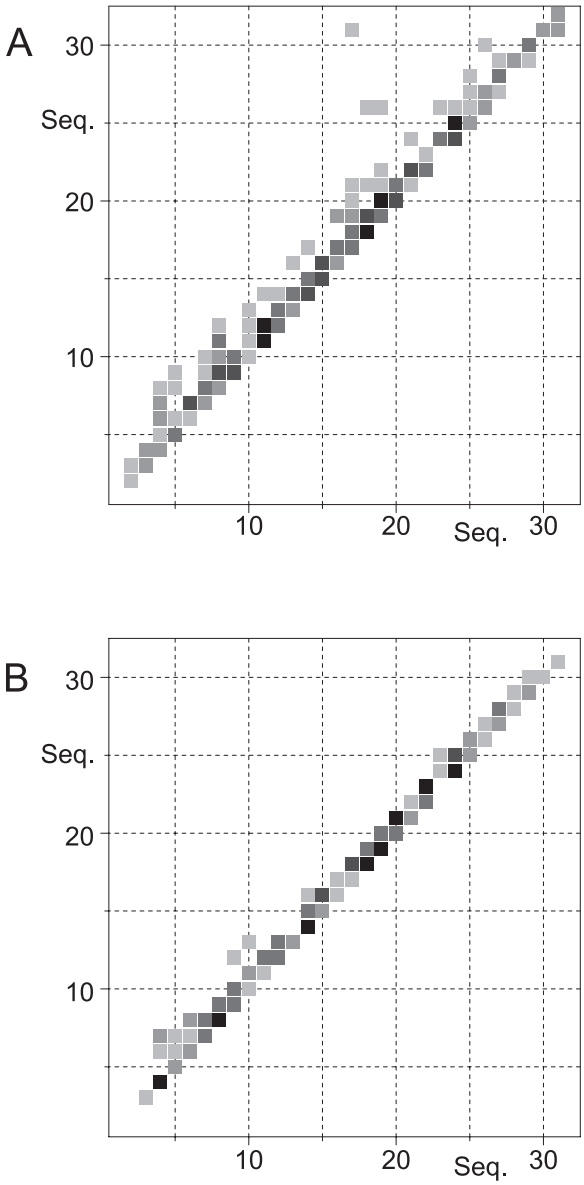


Figure 1. Map of the upper limit distance restraints between residue pairs that were used in the structure calculations of: (A) sCT and (B) hCT. Restraints are classified with different shades according to their number and range from black (number of restraints > 5) to light gray (1 restraint).

with the SANDER module of the AMBER 4.1 package (Weiner and Kollman, 1981; Weiner et al., 1986), using the AMBER91 all-atom parametrization. A distance dependent dielectric constant $\epsilon = r$ was used to roughly take into account solvation effects. A reduction to $\pm$ 0.2 in the net charges of charged side chains was also used to compensate in part for neglect of explicit solvent. A cutoff radius of 0.8 nm for nonbonded

interactions, with a residue-based pairlist routine, was used in all calculations.

The force constants for distance restraints in EM were 500 kJ mol^{-1} nm^{-2}. EM was performed using a combination of steepest descent and conjugate gradient algorithms with a gradient convergence norm of less than 10^{-4} kJ mol^{-1} nm^{-1}.

The best structure of each CT, in terms of energy, underwent 1100 ps of unrestrained MD at constant pressure (1 atm) and constant temperature (300 K), in a periodic rectangular methanol box (2835 solvent molecules, edge lengths of 6.32, 5.96 and 5.92 nm) using the SANDER module of the AMBER 4.1 package (Weiner and Kollman, 1981; Weiner et al., 1986), with the AMBER/OPLS united atom parametrization (Jorgensen and Tirado-Rives, 1988). A dielectric constant $\epsilon = 1$ was used, with a cutoff radius of 0.8 nm for nonbonded interactions, and a residue-based pairlist routine. All covalent bond lengths were constrained to constant values by the SHAKE (Ryckaert et al., 1977) procedure, thus allowing the use of a 2 fs time step. Constant pressure and temperature conditions were obtained by coupling the system with a thermal and a pressure bath (Berendsen et al., 1984), with time constants of 0.1 and 0.5 ps, respectively. The starting solvated system underwent 2000 steps of EM, followed by a 10 ps MD equilibration run. Then the 1100 ps simulation was started, collecting coordinates, velocities and energies every 5 ps. The final statistical analyses were performed on the last 1000 ps.

To account for the inherent flexibility of CT (Castiglione Morelli et al., 1992), unrestrained MD was used to avoid biasing of the structure based on NOE data deriving from multiple conformational equilibria (Jardetzky, 1980; Amodeo et al., 1992). Accordingly, only unambiguous NOEs have been used to validate the sets of final structures, and neither the back-calculations procedure nor the ASNOE routine (Güntert and Wüthrich, 1991) have been applied. The statistical nature of the final structure, represented by a conformational bundle rather than a single conformer, increases the complexity and, simultaneously, reduces the relevance of ambiguous NOEs. Therefore, these data were used neither in the structure generation, nor in the validation procedures.

Molecular structures have been drawn and analyzed with the graphics program MOLMOL (Koradi et al., 1996).

Results and discussion

NMR and distance geometry calculations

Both sCT and hCT have been studied in SDS micelles at pH 4.1 and pH 7.2. Assignment of proton spin systems and secondary structure determination were obtained as previously described (Motta et al., 1991b, 1998). For both calcitonins we observed sequential (αCH_i-NH_{i+1} and NH_i-NH_{i+1}) and medium-range (αCH_i-NH_{i+n}, $1 < n < 4$, and αCH_i-βCH_{i+3}) NOEs but for sCT we confirmed three long-range connectivities ($Thr^{31}NH$-$His^{17}CH_2\beta$, $Asn^{26}NH$-$Lys^{18}CH_2\delta$ and $Asn^{26}NH$-$Leu^{19}CH_\gamma$) (Motta et al., 1991b; Castiglione Morelli et al., 1992). Figure 1 reports the map of the NOEs between residue pairs which were used as upper limit distance restraints in the structure calculations of both sCT and hCT. NOEs are represented by squares between the residues involved, which become darker when the number of NOEs increases. The total number of restraints for sCT (Figure 1A) is larger than that for hCT (Figure 1B) in the region 5–22, while for some residues outside this region (sites 4, 22, 24, 27–29) the situation is reversed.

Three-dimensional structures consistent with NMR data were generated with torsion-angle dynamics carried out with the program DYANA (Güntert et al., 1997). Molecular dynamics in torsion-angle space is claimed to be the preferred structure calculation method for proteins and nucleic acids (Güntert et al., 1997; Stein et al., 1997), except maybe for small helical peptides (Güntert and Mumenthaler, 1997, p. 15). In order to evaluate the possible extension of the method to this class of biomolecules, we performed structure calculations for our peptides using various conditions of simulated annealing (Güntert and Mumenthaler, 1997). In particular, for hCT we calculated 13 sets of conformers, each containing 30 structures, by using different combinations of MD time steps and high temperatures of annealing (namely $N_{steps} = 4000$, 6000, 8000, 10 000 and 20 000, and $T_{high} = 6$, 8, 10, respectively). The best structures in terms of distance restraint violations were obtained either with 8000 or 10 000 MD time steps (Table 1). An increase of the number of time steps above 8000 for the structures calculated at $T_{high} = 6$ (fifth row in Table 1) caused only a small improvement of the average target function value ($Tf = 0.046 \pm 0.016$ nm^2 vs. 0.050 ± 0.022) and of the number of calculated structures having Tf values below 10^{-1} nm^2 (100% vs. 97%). Therefore, whenever the computational ef-

ficiency is an important feature, 8000 steps seem to represent the best choice for this class of peptides. The use of higher temperatures of annealing, as expected, generated a wider distribution of structures though with a general increase of the residual average Tf. However, no appreciable reduction either in the minimum Tf or in the number of structures with Tf $< 10^{-1}$ nm^2 was observed. This suggests that $T_{high} = 6$ already allows an extensive sampling of the most relevant conformations which are compatible with the full set of experimental restraints.

The calculated hCT structures all converged with small residual restraint violations and were further refined by restrained EM using the AMBER program (Weiner and Kollman, 1981). We observed a direct correlation between the values of the DYANA target function and the stability of the refined structures in that the structures with the lowest target functions also had lowest energies.

In order to obtain comparable parameters, 30 sCT structures were calculated at the annealing conditions found to be the best for hCT (8000 MD time steps, $T_{high} = 6$). sCT structures presented no violations of the upper bounds of the NMR distance restraints greater than 0.12 nm, while hCT structures had no violation of the upper bounds greater than 0.10 nm and of the lower bounds greater than 0.12 nm. sCT structures were also subjected to restrained EM: as already found for hCT, the best structures in terms of target function were also the most stable. The structures of both hormones span a similar energy range (-3200 to -3700 kJ mol^{-1} for sCT and -3225 to -3775 kJ mol^{-1} for hCT). In general, refinement produces a reduction of the overall energy but a small increase of the NOE restraint energies is obtained after EM. For sCT, very good convergence was obtained for amino acids in the region Val^8-Lys^{18}. For this region the average root mean square (rms) deviation between structure pairs, using the 12 best structures and including the backbone atoms only, is 0.023 ± 0.009 nm; the corresponding value, including all heavy atoms, is 0.121 ± 0.027 nm. For hCT a unique backbone fold is observed for residues in the region Leu^9-Phe^{16}. The average rms deviation between the 12 best structure pairs is 0.024 ± 0.006 nm, including the backbone atoms in the specified region only, and 0.128 ± 0.026 nm, including all heavy atoms. The distribution of the ϕ and ψ angles (data not shown) of these 12 lowest energy structures shows a reduced dihedral dispersion in the α-helical region for both calcitonins, while the N-terminal residues of both hormones (Cys^1-Ser^5 in sCT and Cys^1-Met^8 in hCT) do not converge to a unique structure. For the C-terminal residues, the progressive angular spreading and the NOE patterns justify the lack of convergence for the regions Pro^{23}-Pro^{32} of sCT, and His^{17}-Pro^{32} of hCT, although the presence of three long-range NOEs in sCT limits the accessible conformational space.

Helix chirality

Among the structures that were not accepted, the 13 worst structures of hCT presented a left-handed α-helix. The chirality of each amino acid was preserved, confirming that use of torsion-angle dynamics excludes inadvertent change of chirality during the course of a structure calculation (Güntert et al., 1997). No inversion of the helix handedness was instead observed for any of the sCT structures. Local or global inversion of the folding of protein structures ('topological mirror images') based on NMR data has already been reported (Pastore et al., 1991; McKnight et al., 1997). As a comparison with the above DYANA structures, 30 structures for hCT and sCT were calculated using the same sets of NOEs as input data, with the DG program DIANA (Güntert et al., 1991), which uses conjugate gradient minimization of a variable target function in torsion-angle space (Braun and Gō, 1985). For hCT, the DIANA structures had target function values and restraint violations higher than those obtained using DYANA (first row in Table 1). The average target function value was 0.082 ± 0.038 nm^2 and the average sum of upper limit NOE restraint violations was 1.08 ± 0.24 nm, with 73% of the calculated structures having Tf values below 10^{-1} nm^2. In all the DYANA runs we always obtained smaller average Tf values (between 0.046 and 0.078) and a greater number of acceptable structures (between 83 and 100%). The values for the DIANA target functions for sCT were found to correspond to those already reported (Motta et al., 1991b). All the hCT and sCT energy-refined structures obtained with DIANA presented a right-handed helix. The fact that DIANA calculations, using the same sets of restraints, do not reproduce the results obtained with DYANA only in the case of hCT suggests that the presence of topological mirror images might be correlated with the different flexibility of the two calcitonins.

For structures derived from NMR spectra free of experimental artifacts, topological mirror images can be observed when:

Table 1. Analysis of NMR restraint violations observed in the hCT structures calculated with DIANA (target function minimization) and with DYANA (simulated annealing with different combinations of T_{high} and N_{steps})

Program[a]	Average target function (Tf) (nm^2)	Minimum Tf (nm^2)	Average $\sum_{\text{upper limit violations}}$ (nm)	Minimum upper limit violation (nm)	Maximum upper limit violation (nm)	Number of structures with Tf $<10^{-1}$ nm^2 (%)
DIANA	0.082 ± 0.038	0.041	1.08 ± 0.24	0.045	0.130	22 (73)
DYANA, 6, 4000	0.057 ± 0.030	0.029	1.00 ± 0.31	0.047	0.101	27 (90)
DYANA, 6, 6000	0.065 ± 0.035	0.033	1.06 ± 0.30	0.046	0.113	26 (87)
DYANA, 6, 8000	0.050 ± 0.022	0.029	0.91 ± 0.22	0.047	0.136	29 (97)
DYANA, 6, 10 000	0.046 ± 0.016	0.026	0.90 ± 0.21	0.036	0.085	30 (100)
DYANA, 6, 20 000	0.060 ± 0.047	0.027	0.95 ± 0.30	0.042	0.136	26 (87)
DYANA, 8, 4000	0.065 ± 0.029	0.031	1.07 ± 0.28	0.044	0.097	26 (87)
DYANA, 8, 6000	0.065 ± 0.039	0.036	1.04 ± 0.31	0.050	0.153	27 (90)
DYANA, 8, 8000	0.061 ± 0.045	0.027	0.99 ± 0.34	0.032	0.137	26 (87)
DYANA, 10, 4000	0.078 ± 0.043	0.035	1.18 ± 0.37	0.047	0.146	25 (83)
DYANA, 10, 6000	0.078 ± 0.052	0.030	1.16 ± 0.42	0.048	0.135	24 (80)
DYANA, 10, 8000	0.064 ± 0.033	0.028	1.07 ± 0.30	0.038	0.098	27 (90)
DYANA, 10, 10 000	0.062 ± 0.025	0.035	1.05 ± 0.29	0.048	0.094	29 (97)
DYANA, 10, 20 000	0.061 ± 0.024	0.030	1.03 ± 0.24	0.047	0.105	28 (93)

[a]The first number refers to the T_{high} value and the second one to N_{steps} (see text for explanation).

(a) an insufficient and not uniformly distributed set of restraints is used for the calculations (undersampling origin, Pastore et al., 1991), or
(b) a highly efficient sampling method, such as torsion-angle MD is applied to a set of distance restraints whose calculated values are averaged by conformational equilibria (averaging origin).

The former origin depends upon the absence of a particular pattern of NOEs which distinguish the two chiralities. In order to uncover which distances can discriminate the chirality in a regular helix from NOE data, we analyzed structural restraints in a right-handed and a left-handed helix. Both forms of an α-helix are conformationally accessible for sequences composed of L amino acids, but steric interactions are estimated to make the left-handed form energetically less favorable by about ca. 8.0 kJ/residue (Ramachandran et al., 1966), and only isolated short segments of this structure are found in proteins (Adzhubei and Sternberg, 1993). Assuming the same absolute values of ϕ and ψ angles for both helices, we found the same NH_i-NH_{i+1} distances but different αCH_i-NH_{i+n} distances for $n = 0, 1, 3, 4$ (Table 2). Thus, discrimination between right-handed and left-handed helix can be achieved by relying upon detectable αCH_i-NH_{i+n} NOE connectivities distributed along the chain. In the present study, the helical segment of hCT was sampled by using a sufficiently high number of restraints containing both NH_i-NH_{i+1} and αCH_i-NH_{i+n} types of connectivities evenly distributed, with more than one 'critical' αCH_i-NH_{i+n} distance per residue. Therefore, the observed chirality inversion of the hCT helix does not depend upon the lack of NOE information (undersampling origin), but is related to averaging of the restraints (origin b).

Table 2. αCH_i-NH_{i+n} distances (nm) in right- and left-handed helix. Identical absolute values of the ϕ (|57°|) and ψ (|47°|) angles for both helices have been assumed. The standard AMBER geometry (Weiner et al., 1984) has been used for bond lengths and valence angles

Helix	n				
	0	1	2	3	4
Right-handed	0.28	0.36	0.45	0.35	0.40
Left-handed	0.23	0.28	0.48	0.47	0.49

For a fast exchanging system, all conformational families contribute the cross-peak intensity thus yielding meaningless virtual conformers if interpreted with a single-conformer model (Jardetzky, 1980). In other words, NOE data of flexible systems cannot be interpreted in terms of a single conformer as the corresponding average interproton distances cannot define a

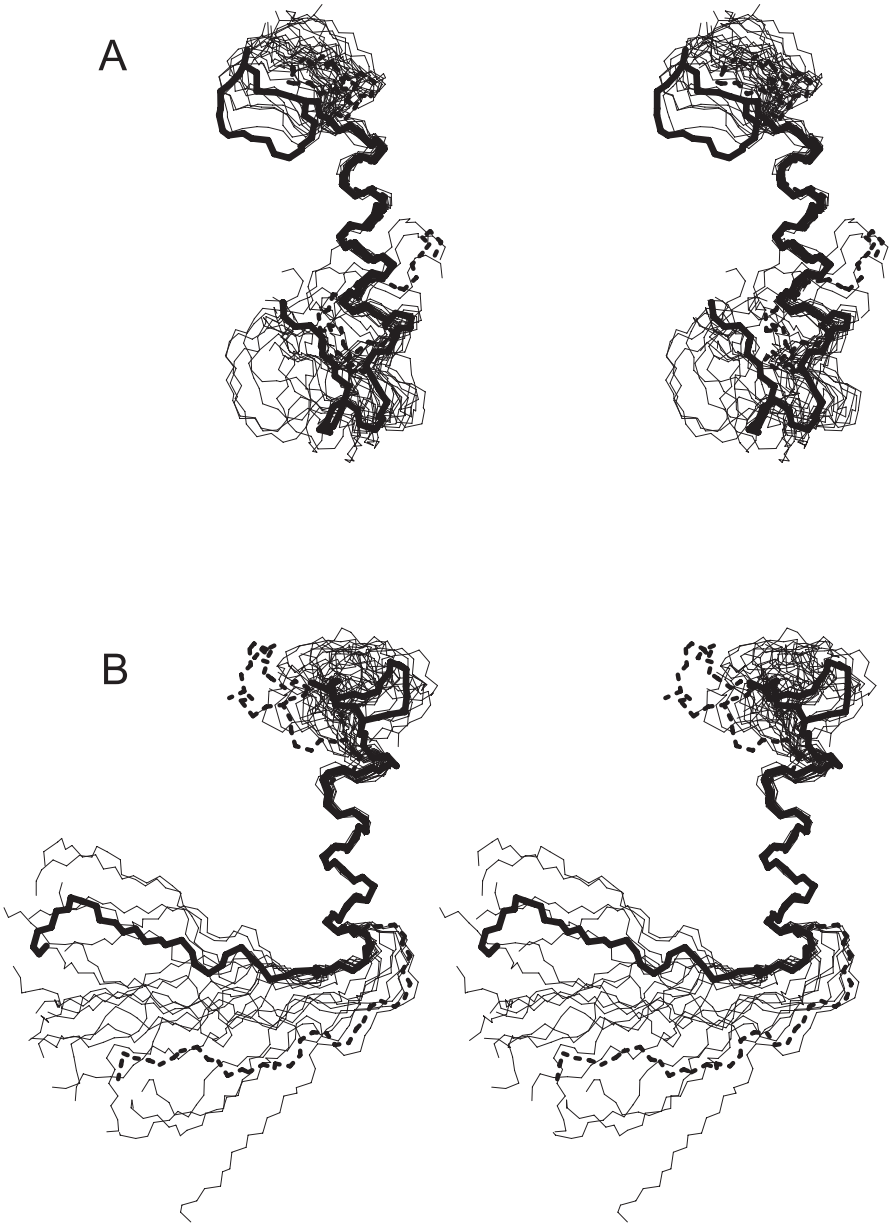


Figure 2. Stereoview of the best fit superposition of 22 frames of the 1100 ps unrestrained MD simulation for (A) sCT and (B) hCT. Structures were sampled every 50 ps. Backbone atoms of residues 13–21 were used for the best fit. Only backbone atoms are shown. The first structure of each run is represented with a dotted line, while a bold line is used for the last.

single structure univocally, and separation of each contribution is not always possible (Amodeo et al., 1992). The increased efficiency of the torsion-angle MD leads to the exploration of a wider region of conformational space, thus allowing the identification of high-energy structures accidentally compatible with the NOE restraints. These physically meaningless structures may suggest either conformational equilibria in solution, or poor-quality cross-peak evaluation.

The number and the distribution of NOEs along the helical segment warrant that hCT is endowed with high conformational flexibility in SDS, and is involved in conformational equilibria. Experimental support comes from CD data which indicate that hCT is more flexible than sCT as hCT adopts the α-helix more

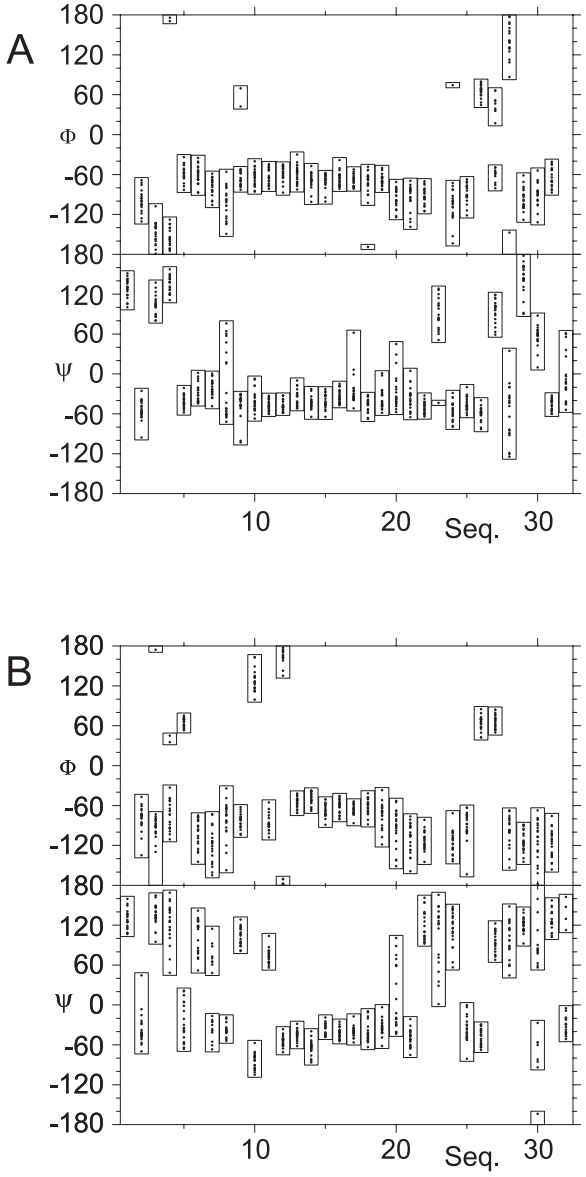


Figure 3. ϕ and ψ distribution by residue of 20 frames of the last 1000 ps of unrestrained MD for (A) sCT and (B) hCT. Structures were sampled every 50 ps. Bars group angle values differing less than 60 degrees from adjacent points.

slowly than sCT (Arvinte and Drake, 1993). Besides, at low temperature sCT adopts an α-helical conformation, whereas hCT showed the occurrence of a positive dichroism, compatible with a non-hydrogen-bonded extended helix (Drake et al., 1988) together with some other unidentified conformations.

MD simulations

The best hCT and sCT structures in terms of potential energy of the EM conformers were subjected to fully unrestrained MD simulations for 1100 ps to better characterize the dynamic properties of both calcitonins. The solvent chosen for the simulation was methanol for the following reasons: (i) methanol is known to favor helix formation for both hCT and sCT with results similar to those of SDS micelles (Epand et al., 1988; Meadows et al., 1991; Arvinte and Drake, 1993); (ii) MD simulations of solvent mixtures are more complex than pure components, especially if one of the solvents is a charged molecule, thus requiring adequate counterions and special (and computationally expensive) care to account for long-range electrostatic interactions; and (iii) the microscopic state of aggregation of SDS in the experimental conditions of NMR spectra is not sufficiently known, and most likely requires too large system sizes, thus reducing considerably the simulated time, and, consequently, the efficiency of conformational sampling for the solute.

The 1100 ps interval turned out to be sufficiently long to observe significant differences in the conformational behavior of both hormones, as also reported for MD simulations on a model helix from the surfactant protein C in different explicit solvents, including methanol (Kovacs et al., 1995). The resulting structures of both calcitonins appear to 'oscillate' in restricted regions of the conformational space, rather than evolving into new completely different conformational states. Figure 2 shows a superposition of 22 snapshots of sCT (Figure 2A) and hCT (Figure 2B) generated every 50 ps of the whole MD simulation. At first glance, the presence of three different conformational domains is evident: a disulphide-bridged N-terminal head, a central helix and a flexible C-terminal tail. The sCT head starts the simulation on the same side of the tail (the starting structure is represented by a broken line in Figure 2A) and at the end of the MD run is always on the same side of the tail although it is rotated of 77 degrees about the helix axis (the last structure is represented by a bold line in Figure 2A). The hCT head starts on the same side of the tail (broken line in Figure 2B) but ends on its opposite side (bold line in Figure 2B). During the simulation the head of both molecules fluctuates between two extreme positions, with a larger spreading observed for hCT. The central helix, although of different length, is the most well defined region in both hormones and it

is characterized by a low mobility. The C-terminal tail of both hormones is the most mobile region. In hCT it is mostly extended, showing no interaction with the helix and spanning an almost conical volume with an angular width of ca. 80 degrees. In sCT, in spite of the fluctuations, the tail is always folded towards the helix forming a cup-like revolution solid.

Statistical analyses of average geometric parameters and of their fluctuations performed in the last 1000 ps of the MD simulation are reported in Figures 3 and 4. The ϕ and ψ distribution of 20 conformations sampled every 50 ps is reported in Figure 3. For sCT (Figure 3A), typical helical values were found in the region Ser^{5}-Tyr^{22}, with rms deviations from 9 to 17 degrees with the exception of residues Val^{8}, Leu^{9}, His^{17} and Gln^{20}, for which higher values were found. Residues Leu^{9}-Leu^{19} are constantly detected in a helical conformation during the whole trajectory, although average ϕ and ψ values in the Ser^{5}-Val^{8} and Gln^{20}-Tyr^{22} ranges are similar to those detected in the 9–19 helix. For hCT (Figure 3B) a larger spreading is observed for almost all the hCT sequence, but characteristic helical values were found in the region Thr^{13}-Thr^{21} with rms deviations from 8 to 19 degrees (with higher values for residues His^{20} and Thr^{21}). In addition, the constantly helical region is reduced to the Thr^{13}-Phe^{19} region, with residues outside this range exhibiting markedly non-helical average values.

The real-time variations of ϕ and ψ dihedral angles, taken from snapshots between 100 and 1100 ps at 50 ps intervals during the MD trajectory, are presented in Figure 4. Visual inspection of these plots confirms the general stability of the sCT helix in the region 5–22 (Figure 4A). Only very localized correlated variations are observed for ψ_8 and ϕ_9, and, outside the helical region, for ψ_{23} and ϕ_{24}, for ψ_{27} and ϕ_{28}, and ψ_{28} and ϕ_{29}. In hCT small variations of the ϕ and ψ angles are observed for helical residues 13–19 while correlated variations are observed in wider regions at the N-terminus from ψ_6 to ϕ_8 and at the C-terminus from ψ_{22} to ψ_{25} and from ψ_{27} to ϕ_{31} (Figure 4B). An important feature of the sCT C-terminal region is the absence of large fluctuations up to the last residue, which confirms the influence of long range helix-tail interactions in the stabilization of the overall peptide folding.

A more quantitative representation of the dynamic properties of calcitonins is reported in Figure 5 which compares the ϕ and ψ angular order parameter (S) values (Hyberts et al., 1992) for sCT (Figure 5A) and hCT (Figure 5B). A high order with $S > 0.9$ is observed in the N- and C-termini of sCT, as well as in the central region; low-order values ($S < 0.75$) are localized to residues 8, 23, 24, 27 and 28. These data quantitatively support the 'factorization' of sCT into three different conformational regions (Amodeo et al., 1994) separated by flexible residue(s): Val^{8} connecting the head and the helix, and Pro^{23}-Arg^{24} connecting the helix and the tail. In hCT, amino acids with $S < 0.9$ are spread along the whole chain (residues 4, 7, 8, 11, 12, 20, 21, 23, 24, 28, 30 and 32), indicating a distributed flexibility. An analysis of the NMR restraint violations for the 20 conformations of calcitonins, sampled every 50 ps in the last 1000 ps of the unrestrained MD run, is shown in Table 3. The selected structures represent a statistically meaningful ensemble for the analysis of both peptides, as indicated by the absence of systematic violations of experimental restraints. The number of violations has been normalized so that the results for the two calcitonins are comparable, in spite of the different number of experimental restraints. The number of violations larger than 0.02 nm is higher in hCT (40%) than in sCT (33%), while the number of violations > 0.15 nm is slightly higher for sCT. The observed hydrogen bonding pattern confirms that the start of the helix is shifted towards the C-terminus during the MD run, whereas hydrogen bonds in sCT are slightly violated only in 5 structures out of 20. Besides, for sCT the largest violations were found in NOEs involving side-chain protons. This analysis, as well as the other structural features discussed for the restrained conformations, shows substantial differences in the conformational behavior of the two calcitonins, affecting both the nature and the amount of their conformational freedom. In fact, hCT exhibits a higher overall flexibility, uniformly distributed along

Table 3. Analysis of NMR restraint violations observed in 20 conformations of calcitonins generated every 50 ps during the last 1000 ps of the unrestrained MD run. The number of violations has been normalized, dividing it by the total number of restraints (all), and is reported as percentage. In the third and fifth columns side-chain NOEs are not included (no SC)

NOE violation (nm)	hCT (%)		sCT (%)	
	All	No SC	All	No SC
> 0.02	40	39	33	30
> 0.05	22	21	26	24
> 0.10	14	13	16	13
> 0.15	7	7	11	9

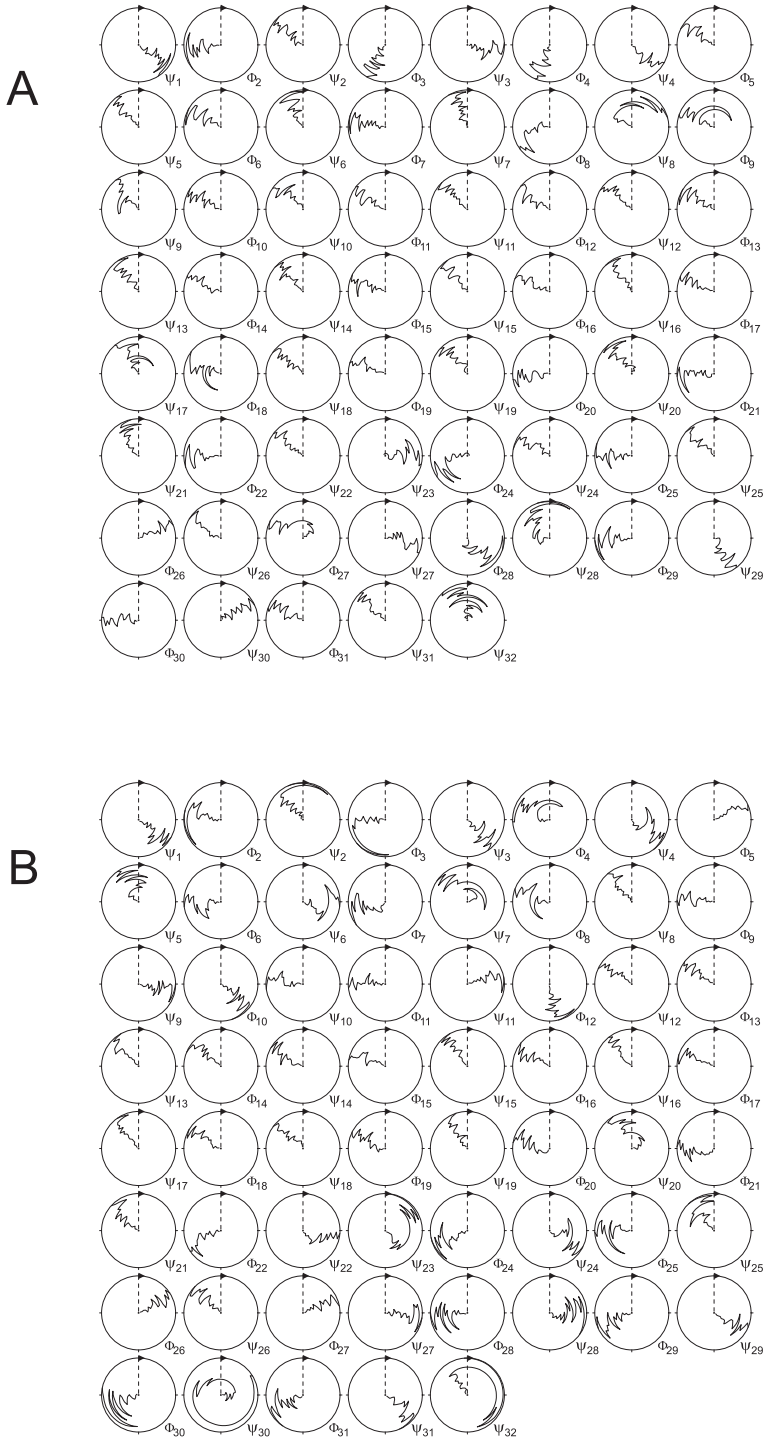


Figure 4. Real-time variations of the ϕ and ψ angles through the last 1000 ps of unrestrained MD for (A) sCT and (B) hCT. Snapshots were taken at 50 ps intervals.

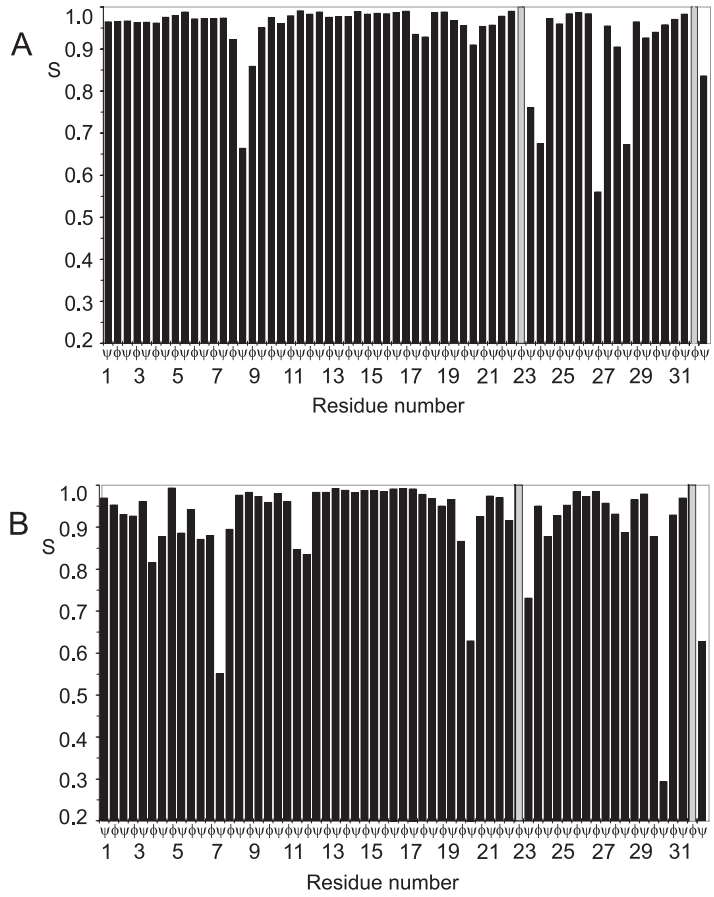


Figure 5. Angle order parameters S(ϕ) and S(ψ) by residue in the last 1000 ps of unrestrained MD for (A) sCT and (B) hCT. S(ϕ) of Pro residues is shown in light gray.

the polypeptide chain and associated to a rather continuous distribution of similar conformers. This accounts for the simultaneous more-small-violation vs. less-large-violation trend observed in comparison with sCT in a 1-ns unrestrained MD. On the contrary, the conformational behavior of sCT seems to be dominated by a relatively small number of well-clustered conformers, exhibiting large differences in the folding of their C-terminal domains and a limited freedom in the central helix. These features are completely independent from the choice and application of experimental restraints, as they emerge both in restrained and in unrestrained calculations.

Conclusions

The predominant conformational feature of both calcitonins is an amphipathic α-helix. For sCT, the ϕ and ψ distribution indicates that residues in the region Leu^{9}-Leu^{19} are constantly detected in a helical conformation during the whole trajectory of the 1100 ps MD simulations, although angular values for boundary amino acids are close to those for helical region. In hCT, a shorter helical segment is found in Thr^{13}-Phe^{19} region, but average characteristic angular helical values were found in the Thr^{13}-Thr^{21} region. An analysis of the NMR restraints violations together with the angular order parameters S(ϕ) and S(ψ) indicate that sCT presents very localized stretches of 'flexible' amino acids connecting the head and the helix, and the helix and the tail. On the contrary, hCT exhibits a higher flexibility, more uniformly distributed along the polypeptide chain, including the helix. It is noteworthy that DG in torsion-angle space obtained with the DYANA algorithm uncovers a left-handed α-helix for hCT but not for any of the sCT structures. Since the number of restraints in both calcitonins is sufficiently high and well distributed along

the polypeptide chain, the presence of a left-handed α-helix stems from averaged NOE restraints, which in turn depend upon the flexibility of the helical region in hCT. In fact, restraints in flexible systems cannot be interpreted in terms of a single conformer as the corresponding average interproton distances cannot define a single structure univocally. On the other hand, it is known from CD studies that hCT adopts the α-helix more slowly than sCT (Arvinte and Drake, 1993), and that at low temperatures hCT showed the occurrence of a positive dichroism, indicating the presence of a polyproline-II helix (Drake et al., 1988). Taken together, our results indicate that hCT structures exist in a continuous distribution of similar interconverting conformers endowed with high mobility; sCT structures are clustered in small numbers of families, and each structure is 'factorized' into three different conformational regions with a limited freedom for the central helix.

The torsion-angle dynamics method turned out to be efficiently applicable to the structure calculation of small helical polypeptides, provided that a careful choice of the length and of the temperature of the simulated annealing is made. An interesting finding is that the accuracy of the sampling is able to uncover conformational equilibria in solution by detecting spurious high-energy structures, which for right-handed α-helices are likely to be represented by a left-handed α-helix with all the amino acids in the L configuration.

References

Adzhubei, A.A. and Sternberg, M.J.E. (1993) *J. Mol. Biol.*, **229**, 472–493.

Amodeo, P., Motta, A., Tancredi, T., Salvadori, S., Tomatis, R., Picone, D., Saviano, G. and Temussi, P.A. (1992) *Pept. Res.*, **5**, 48–55.

Amodeo, P., Castiglione Morelli, M.A. and Motta, A. (1994) *Biochemistry*, **33**, 10754–10762.

Arvinte, T. and Drake, A.F. (1993) *J. Biol. Chem.*, **268**, 6408–6414.

Austin, L.A. and Heath III, H. (1981) *N. Engl. J. Med.*, **304**, 269–278.

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

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

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A. and Haak, J.M. (1984) *J. Chem. Phys.*, **81**, 3684–3690.

Braun, W. and Gō, N. (1985) *J. Mol. Biol.*, **186**, 611–626.

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

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

Dayhoff, M.O. (1978) *Atlas of Protein Sequence and Structure*, Vol. 5, Suppl. 3, National Biomedical Research Foundation, Silver Spring, MD, USA, p. 149.

Doi, M., Kobayashi, Y., Kyogoku, Y., Takimoto, M. and Goda, K. (1990) In *Peptides: Chemistry, Structure and Biology, Proceedings of the Eleventh American Peptide Symposium* (Rivier, J.E. and Marshall, R., Eds), ESCOM, Leiden, The Netherlands, pp. 165–167.

Drake, A.F., Siligardi, G. and Gibbons, W.A. (1988) *Biophys. Chem.*, **31**, 143–146.

Drobny, G., Pines, A., Sinton, S., Weitekamp, D. and Wemmer, D. (1979) *Faraday Symp. Chem. Soc.*, **13**, 49–55.

Epand, R.M., Epand, R.F., Orlowski, R.C., Schlueter, R.J., Boni, L.T. and Hui, S.W. (1983) *Biochemistry*, **22**, 5074–5084.

Epand, R.M., Epand, R.F., Orlowski, R.C., Seyler, J.K. and Colescott, R.L. (1986) *Biochemistry*, **25**, 1964–1968.

Epand, R.M., Epand, R.F. and Orlowski, R.C. (1988) *Biochem. Biophys. Res. Commun.*, **152**, 203–207.

Green III, F.R., Lynch, B. and Kaiser, E.T. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8340–8344.

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

Güntert, P., Braun, W., Billeter, M. and Wüthrich, K. (1989) *J. Am. Chem. Soc.*, **111**, 3997–4004.

Güntert, P. and Wüthrich, K. (1991) *J. Biomol. NMR*, **1**, 447–456.

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

Güntert, P. and Mumenthaler, C. (1997) *DYANA Manual, version 1.4*, Institut für Molekularbiologie und Biophysik, ETH, CH-8093 Zürich, Switzerland.

Güntert, P., Mumenthaler, C. and Wüthrich, K. (1997) *J. Mol. Biol.*, **273**, 283–288.

Hyberts, S.G., Goldberg, M.S., Havel, T.F. and Wagner, G. (1992) *Protein Sci.*, **1**, 736–751.

Jardetzky, O. (1980) *Biochim. Biophys. Acta*, **621**, 227–232.

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

Jorgensen, W.L. and Tirado-Rives, J. (1988) *J. Am. Chem. Soc.*, **110**, 1657–1664.

Koradi, R., Billeter, M. and Wüthrich, K. (1996) *J. Mol. Graph.*, **14**, 51–55.

Kovacs, H., Mark, A., Johansson, J. and van Gunsteren, W. (1995) *J. Mol. Biol.*, **247**, 808–822.

Levy, F., Muff, R., Dotti-Sigrist, S., Dambacher, M.A. and Fisher, J.A. (1988) *J. Clin. Endocrinol. Metabol.*, **67**, 541–548.

McKnight, C.J., Matsudaira, P.T. and Kim, P.S. (1997) *Nat. Struct. Biol.*, **4**, 180–185.

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

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

Moe, G.R. and Kaiser, E.T. (1985) *Biochemistry*, **24**, 1971–1975.

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

Motta, A., Castiglione Morelli, M.A., Goud, N.A. and Temussi, P.A. (1989) *Biochemistry*, **28**, 7996–8002.

Motta, A., Temussi, P.A., Wünsch, E. and Bovermann, G. (1991a) *Biochemistry*, **30**, 2364–2371.

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

Motta, A., Andreotti, G., Amodeo, P., Strazzullo, G. and Castiglione Morelli, M.A. (1998) *Proteins*, **32**, 314–323.

Neuhaus, D., Wagner, G., Vasák, M., Kägi, J.H.R. and Wüthrich, K. (1985) *Eur. J. Biochem.*, **151**, 257–273.

Neidig, K.P., Geyer, M., Goerler, A., Antz, C., Saffrich, R., Beneicke, W. and Kalbitzer, H.R. (1995) *J. Biomol. NMR*, **6**, 255–270.

Pastore, A., Atkinson, R.A., Saudek, A. and Williams, R.J.P. (1991) *Proteins*, **10**, 22–32.
Piantini, U., Sørensen, O.W. and Ernst, R.R. (1982) *J. Am. Chem. Soc.*, **104**, 6800–6801.
Piotto, M., Saudek, V. and Sklenar, V. (1992) *J. Biomol. NMR*, **2**, 661–666.
Ramachandran, G.N., Venkatachalam, C.M. and Krimm, S. (1966) *Biophys. J.*, **6**, 849–872.
Ryckaert, J.P., Ciccotti, G. and Berendsen, H.J.C. (1977) *J. Comput. Phys.*, **23**, 327–341.
Stein, E.G., Rice, L.M. and Brünger, A.T. (1997) *J. Magn. Reson*., **124**, 154–164.
Wagner, G., Braun, W., Havel, T.F., Schaumann, T., Gō, N. and Wüthrich, K. (1987) *J. Mol. Biol.*, **196**, 611–639.
Weiner, P.K. and Kollman, P.A. (1981) *J. Comput Chem.*, **2**, 287–303.
Weiner, S.J., Kollman, P.A., Case, D.A., Chandra Singh, U., Ghio, C., Alagona, G., Profeta, S. and Weiner, P.K. (1984) *J. Am. Chem. Soc.*, **106**, 765–784.
Weiner, P.K., Kollman, P.A., Nguyen, D.T. and Case, D.A. (1986) *J. Comput. Chem.*, **7**, 230–252.
Wüthrich, K., Billeter, M. and Braun, W. (1983) *J. Mol. Biol*., **169**, 949–961.