

Rational design of aggregation-resistant bioactive peptides: Reengineering human calcitonin

Susan B. Fowler*[†], Stephen Poon*[†], Roman Muff^{†‡}, Fabrizio Chiti[§], Christopher M. Dobson*[¶], and Jesús Zurdo*^{¶||}

*Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom; [†]Research Laboratory for Calcium Metabolism, Departments of Orthopedic Surgery and Medicine, University of Zurich, Balgrist University Hospital, Forchstrasse 340, 8008 Zurich, Switzerland; and [§]Dipartimento di Scienze Biochimiche, Università di Firenze, Viale Morgagni 50, 50134 Florence, Italy

Edited by James A. Wells, Sunesis Pharmaceuticals, Inc., South San Francisco, CA, and approved May 23, 2005 (received for review February 11, 2005)

A high propensity to aggregate into intractable deposits is a common problem limiting the production and use of many peptides and proteins in a wide range of biotechnological and pharmaceutical applications. Many therapeutic polypeptides are frequently abandoned at an early stage in their development because of problems with stability and aggregation. It has been shown recently that parameters describing the physicochemical properties of polypeptides can be used as predictors of protein aggregation. Here we demonstrate that these and similar tools can be applied to the rational redesign of bioactive molecules with a significantly reduced aggregation propensity without loss of physiological activity. This strategy has been exemplified by designing variants of the hormone calcitonin that show a significantly reduced aggregation propensity, yet maintain, or even increase, their potency when compared to the current therapeutic forms. The results suggest that this approach could be used successfully to enhance the solubility and efficacy of a wide range of other peptide and protein therapeutics.

protein aggregation | protein design | biopharmaceuticals | amyloid | misfolding

The number of pharmacologically active peptides and proteins under development for the prevention and treatment of human disorders is increasing, and as a result, so is the pressure to overcome problems associated with aggregation and stability. Up to 96% of all drug candidates in trials are abandoned during preclinical or clinical development, often because of low solubility or aggregation problems (1). Many peptide-based drugs with great therapeutic potential are rendered ineffective simply because of an intrinsic propensity to aggregate irreversibly (2). Aggregation is one of the most significant obstacles to the development of protein-based drugs because it cannot only compromise their bioavailability and therapeutic activity but it may also increase the risk of immunogenic reactions (3, 4).

A good example of a bioactive peptide with limited pharmaceutical potential due to a high tendency to aggregate is human calcitonin (hCT), a 32-residue polypeptide hormone synthesized and secreted by the C cells of the thyroid, and involved in calcium regulation and bone dynamics. *In vivo*, calcitonin (CT) causes a rapid, but short-lived, decline in calcium and phosphate levels in the blood by promoting the incorporation of these ions into bone (5). This activity has led to the use of CT in the treatment of conditions such as osteoporosis and Paget's disease, as well as malignancy-caused hypercalcemia and musculoskeletal pain (5–7). However, hCT shows an extremely high tendency to self-associate in the form of amyloid fibrils. As a result hCT amyloid deposits can occur *in vivo* in patients with medullar carcinoma of the thyroid (8, 9) and *in vitro* in preparations designed for patient administration (10). Not only does aggregation constitute a serious problem during the production, storage, and administration of hCT, but it can also lead to a significant decrease in its activity as a drug (11). Moreover, several studies in patients suggest that aggregation stimulates undesirable immune responses resulting in resistance or allergic reactions (12–14).

Some cell-based studies indicate that aggregation also increases drug-induced cytotoxicity in the case of CT as well as other therapeutics (15, 16).

Despite its low sequence identity with hCT (50%), salmon calcitonin (sCT) has been the clinically preferred alternative to hCT for several years because of its substantially lower propensity to aggregate (5). However, the use of sCT has been shown to cause side effects, such as anorexia and vomiting (17, 18), as well as to trigger immune reactions that could severely limit its effectiveness as a drug (19–23). Interestingly, it has been reported that hCT shows a much higher potency than the salmon variant under specific conditions where aggregation is prevented (24). Such specific conditions are difficult to implement during the production, storage, and distribution of hCT and its administration to patients, explaining why hCT has never been extensively used as a therapeutic. These findings suggest that a human-like CT variant with a reduced aggregation propensity, but retaining functionality, could represent a valuable therapeutic alternative to sCT.

One strategy for overcoming these problems involves the development and screening of specifically designed analogues that, incorporating just a small number of amino acid substitutions, show normal or improved biological activity as well as reduced aggregation propensity, with respect to their wild-type predecessors. Recently, it has been shown that a combination of simple physicochemical parameters can be used to predict, in a quantitative manner, the effect of single mutations on the aggregation behavior of a range of polypeptides (25, 26). In the light of the success of this approach in predicting the effect of single amino acid substitutions associated with protein deposition diseases (26), we wished to expand the applications of this method (*i*) to determine whether a similar combination of parameters defining the properties of a polypeptide chain could be used to predict the aggregation behavior of sequences with multiple amino acid substitutions; and (*ii*) to assess whether this approach could be used successfully to design rationally new polypeptide variants with reduced aggregation propensities. Here, we demonstrate the success of this approach by showing that full-length hCT-like variants containing just a small number of rationally designed amino acid substitutions show greatly improved solubility yet maintain their physiological activity.

Materials and Methods

Aggregation Studies. CT_{10–21} peptides (Southampton Polypeptides, Southampton, U.K.) and full-length CT (Biopeptide, San Diego) peptides were synthesized by using standard Fmoc synthesis. CT_{10–21} peptides were N-terminally acetylated and C-terminally amidated and were supplied with a purity >95%.

This paper was submitted directly (Track II) to the PNAS office.

Freely available online through the PNAS open access option.

Abbreviations: CT, calcitonin; hCT, human CT; sCT, salmon CT.

[†]S.B.F., S.P., and R.M. contributed equally to this work.

[¶]C.M.D. and J.Z. are cofounders of Zyentia, Ltd.

^{||}To whom correspondence should be addressed. E-mail: jz226.cam.ac.uk.

© 2005 by The National Academy of Sciences of the USA

Full-length CT peptides were C-terminally amidated (essential for bioactivity) and supplied as acetate salts with a purity >98%. Details on sample preparation and kinetic measurements are included in the supporting information, which is published on the PNAS web site.

The rate of aggregation in the exponential growth phase was determined by fitting the data to a single exponential function plus a term to account for baseline drift ($Y = Ae^{-kx} + Bx + C$). Aggregation propensity for all of the sequences analyzed was estimated by using algorithms described elsewhere (26–28) (accessible online at www.amyloidfibril.com) as described in the supporting information.

Thioflavin T binding and negative electron microscopy were performed at the end of each aggregation experiment to confirm the presence of amyloid fibrils as reported (29–31) (see supporting information).

Cell Culture and Biological Activity Assays. T47D cells were grown in Dulbecco's modified Eagle's medium containing 4.5 g/liter glucose and Ham F12 medium (1:1) supplemented with 2 mM L-glutamine, 100 nM dexamethasone, and 10% FCS in 5% CO₂, as described (32). For binding studies and cAMP measurements the cells were seeded into 24-well plates at a density of 100,000 cells per well and allowed to reach confluence. Lyophilized peptide aliquots were dissolved in PBS containing 20 mg/ml D-mannitol to a stock concentration of 100 μ M and used immediately. Fresh stocks were prepared for each experiment. hCT was iodinated with the use of a modified chloramine T method and purified by HPLC as described (32–34). CT receptor binding, cAMP stimulation, and hypocalcaemic activity assays were performed as described in the supporting information.

Results and Discussion

Aggregation Studies on CT_{10–21} Peptides. To test the ability to predict the effect of multiple amino acid substitutions on the aggregation rate of a polypeptide sequence, a series of short peptides were designed by using amino acids 10–21 of hCT as an initial template (these peptides are collectively named CT_{10–21}). Studies on peptide fragments of hCT have suggested that residues in this region could play an important role in the activity and aggregation behavior of the full-length peptide (35, 36). It has been shown in particular that five residues corresponding to positions 15–19 of hCT (DFNKF) play an active role in oligomerisation and fibril formation by hCT *in vitro* (35), and Lys-18 and Phe-19 have been identified as key residues in both the bioactivity and self-assembly of hCT (36).

Sequence variability between CT_{10–21} peptides involve amino acid substitutions designed to modify a wide range of parameters including the net charge, intrinsic hydrophobicity, and secondary structure propensities (see supporting information). The aggregation rate constants for each of the CT_{10–21} peptides were calculated from kinetic experiments monitored by changes in turbidity at 340 nm over time (Fig. 1*a*). Full details of these experiments can be found in the supporting information.

Of the 17 CT_{10–21} peptide variants tested, six did not show any measurable increase in turbidity in the timeframe of the experiment (>60 days) in any of the different conditions tested (supporting information). However, the remaining 11 samples showed increasing turbidity as a function of time, and TEM analysis confirmed the presence of amyloid fibrils in these preparations. Fig. 1*a* shows a plot of the relative aggregation rates for the 11 CT_{10–21} peptide variants and includes reference data for single point mutants of natively unfolded polypeptides associated with amyloid depositional disorders, including IAPP, the amyloid β -peptide and α -synuclein (26). The combined plot has a linear correlation coefficient of $r = 0.87$ ($P < 0.02$) and a slope of 1.01. The correlation between the predicted and experimental values for the CT_{10–21} variants is highly significant, showing that the empirical algorithm

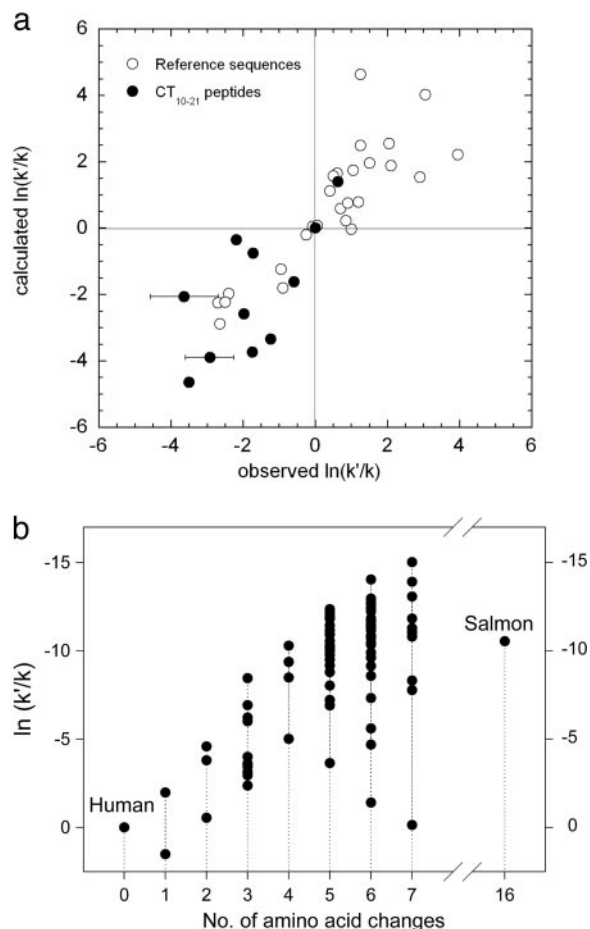


Fig. 1. Prediction of the aggregation rates of CT_{10–21} peptides. (a) Comparison of the experimental relative aggregation rates for hCT_{10–21} peptides with predicted values calculated as reported (26). Relative aggregation rates are plotted as $\ln(k'/k)$, where k is the rate of the human wild-type sequence and k' that of the mutant sequence. Filled circles correspond to the values obtained for those CT_{10–21} peptides that were amenable to measurement in the time span of the experiment. Empty circles correspond to a data set of mutants of a range of natively unfolded polypeptides (amylin, A β -peptide, tau, α -synuclein, prion peptides) as described (26). (b) *In silico* aggregation propensity predictions for a collection of CT variants with an increasing number of modifications with respect to full-length hCT. The predicted value for sCT (16 modifications) is included for comparison. Details of the sequences are included in supporting information.

can be applied successfully to predict the changes in aggregation rates resulting from multiple as well as just single substitutions. At the end of the experiment, the amount of peptide remaining in solution was also measured for all samples. Interestingly, the amount of peptide recovered shows a significant degree of correlation with the predicted aggregation propensity (supporting information). Altogether, the data support the use of this algorithm as a valid strategy for redesigning hCT, and, ultimately, other polypeptides, to increase substantially their stability in solution against aggregation.

Redesigning Full-Length CT Variants. In the light of these findings, a series of full-length variant CT sequences was designed on the basis of the following principles. First, modifications to residues thought to be important for physiological function were avoided to minimize the risk of compromising the biological activities of the variants, and hence their therapeutic potential. The first seven residues of hCT (CGNLSTC), which are believed to be important in the activation of the CT receptor (37), were left

a

Peptide	Sequence
Human	CGNLSTCMLGTYTQDFNKFHTFPQTAIGVGAP-NH ₂
Salmon	C SNLSTCVL GKLSQELHKLQTYPRNT TGSGTP-NH ₂
6CHR	CGNLSTCML GKLTQELRKLHTFPQTAIGVGAP -NH ₂
5P	CGNLSTCML GRYTQDFRKFHTFPRTATGSGAP -NH ₂
6S	CGNLSTCML GRYTQDFRKFRTFPRTATGSGAP -NH ₂

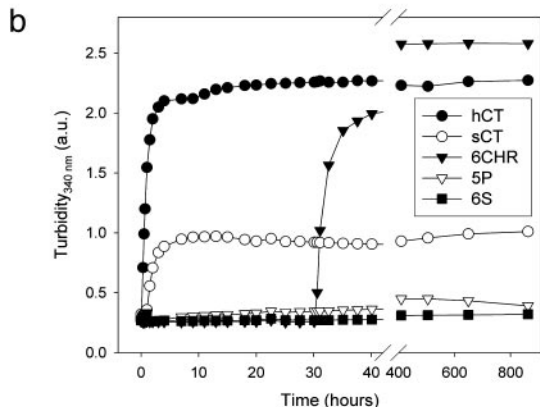


Fig. 2. Aggregation kinetics of different full-length CT variants. (a) Amino acid sequences of hCT, sCT, and the designed variants 6CHR, 5P, and 6S. Amino acids that differ from the human sequence are marked in bold. (b) Aggregation kinetics for all five calcitonin variants in PBS, pH 7.4 at 37°C.

unchanged in all redesigned variants. Second, modifications that are predicted to increase α -helical stability were introduced in the area of the peptide with the highest helical propensity in the wild type sequence (residues 13–19). This approach was followed based on evidence obtained in other systems indicating that stabilization of native-like local helical interactions could have a beneficial effect in preventing aggregation (30, 38, 39). Third, the maximum allowed number of changes from the wild-type sequence was kept to 20% of the total number of residues, with the intention of maintaining biological activity and minimizing potential side effects. Modifications were also made with the specific aims of reducing the net hydrophobicity of the polypeptide chain and disrupting hydrophobic patches present in hCT, especially those located in highly aggregation-prone regions of the molecule. We used a modification of the algorithms reported in refs. 26 and 27 to identify highly aggregation-prone regions

(aggregation hotspots) in the CT molecule (28) (supporting information).

The aggregation propensities of >600 sequences were analyzed *in silico* by using the algorithm tested on the short peptide sequences (26). Fig. 1b shows the results of this analysis on a subset of those sequences (see supporting information for details). There is a general decrease in aggregation propensity as a larger number of mutations are incorporated into the sequence of hCT, whereas the spread of values shows the sensitivity of the aggregation rate to the specific nature of the substitution. Remarkably, four modifications are predicted to be sufficient to reduce the aggregation propensity of hCT to that of sCT, the currently favored therapeutic. This result supports the rationale that a small number of substitutions could be sufficient to generate a version of hCT able to surpass the stability of sCT against aggregation.

Aggregation Kinetics of Redesigned Variants. Three full-length CT variant sequences were selected for testing in experimental assays (see Fig. 2A). Two of these variants (5P and 6S) were chosen because of their high scores in the *in silico* selection procedure illustrated in Fig. 1b. In both cases, only those residues less conserved across CT sequences of different species were varied. When several modifications gave similar scores, selection was performed by giving priority first to residues present in hCT, followed by those present in CT sequences found in other species. Otherwise, the most conservative mutations were chosen (e.g., Lys to Arg). The third variant (6CHR) was designed with the specific aim of substantially enhancing the helical propensity of the peptide, because it has been shown that stabilization of local helical interactions can reduce the aggregation propensity of peptides and partially unfolded proteins (30, 38, 39). Finally, the maximum number of mutations included was chosen to be six to keep sequence identity with the parental hCT to a value >80%.

The time courses of aggregation of the three variants are shown in Fig. 2b together with those of hCT and sCT (see also Table 1). A significant reduction in the rate of aggregation can be observed for all reengineered variants when compared with the human sequence. This is particularly dramatic for variants 5P and 6S, which not only show a substantially decreased aggregation propensity compared to hCT, but also improve that of sCT. The success of the present design strategy is not only evident in the reduced rates of aggregation of the engineered variants, as revealed by turbidometry, but also in the amount of material remaining in solution at the end of the aggregation experiment. Amino acid analysis showed that $\approx 32\%$ and 74% of the starting material for variants 5P and 6S, respectively, was still present in solution at the end of the incubation

Table 1. Kinetics of aggregation of full-length natural and designed CT variants

Peptide	Helical content*	Predicted $\ln(k'/k)^{\dagger}$	Observed $\ln(k'/k)$	Predicted $\ln(k_{agg}), s^{-1*}$	Observed $\ln(k_{agg}), s^{-1}$	$\ln(T_{lag}), s^{\S}$	ΔA_{340}	% recovery
hCT	0.17	0	0	-10.2	-7.6 ± 0.2	6.9 ± 0.2	1.9 ± 0.3	0.7 ± 0.2
sCT	0.75	-10.5	-0.3	-12.6	-7.9 ± 0.3	8.3 ± 0.1	0.8 ± 0.2	7.0 ± 2.0
6CHR	2.19	-5.6	-1.1	-11.3	-8.7 ± 0.3	11.6 ± 0.7	2.4 ± 0.3	0.8 ± 0.2
5P	0.63	-11.4	-2.8	-12.8	-10.4 ± 1.4	9.9 ± 0.7	0.4 ± 0.2	32.0 ± 6.0
6S	0.69	-14.0	-4.5	-13.5	-12.2 ± 0.9	11.3 ± 1.2	0.2 ± 0.1	74.0 ± 5.0

Each measurement presented is the mean of at least four different experiments performed on different days, and each consisting of five replicates, along with the standard deviation. Aggregation was also monitored by thioflavin-T binding and TEM analysis (supporting information).

*Intrinsic helical propensity (in %) was calculated with AGADIR (48) using the appropriate experimental conditions.

[†]Calculated as described (26).

[‡]Calculated using an extended version of the algorithm described in Chiti *et al.* (26) as reported elsewhere (27, 28), including sequence patterns that favour aggregation (49), and predicts absolute aggregation rates for any sequence from a random-coil state (www.amyloidfibril.com).

[§] T_{lag} is a descriptor of the lag phase, and corresponds to the time taken to reach 10% of the final amplitude of each kinetic trace.

^{||}Recovery was estimated by amino acid analysis of the different samples after sedimentation at $13,000 \times g$ and filtering through a $0.2\text{-}\mu\text{m}$ filter.

Table 2. Values of [¹²⁵I]hCT binding inhibition and cAMP stimulation

	[¹²⁵ I]CT binding inhibition IC ₅₀ , pM	cAMP stimulation EC ₅₀ , pM
hCT	3220 ± 240	231 ± 73*
sCT	224 ± 25 [†]	131 ± 18**
6CH-R	416 ± 40 [†]	96 ± 20**
6S	408 ± 106 [†]	115 ± 27**
5P	154 ± 34 [†]	48 ± 7

Results are means ± SEM. Binding and cAMP experiments were repeated three and six times, respectively. *, $P < 0.05$ versus 5P. †, $P < 0.01$ versus hCT. ‡, $P < 0.05$ versus hCT when treated as a group of peptides of equivalent potency.

period (60 days), compared to 7% for sCT and <1% for hCT (Table 1). This enhanced solubility has important benefits for the long-term storage and bioavailability of these variants, as well as a decreased risk of undesired immunogenic responses associated with aggregation (12–14).

Physiological Activity of hCT Variants. The biological activities of the redesigned CT variants were tested by monitoring their ability to

recognize and activate specifically the hCT receptor in T47D (human ductal carcinoma) cells. This cell line primarily expresses the receptor hCT_(a) (formerly known as hCRT2) as well as lesser amounts of receptor hCT_(b) (formerly known as hCRT1), in a similar expression pattern to that observed in osteoclasts, the ultimate target for hCT in the treatment of osteoporosis, especially in postmenopausal women (40, 41). Receptor binding for each of the full-length hCT variants was monitored by a competition assay under equilibrium conditions employing [¹²⁵I]hCT. Total [¹²⁵I]hCT binding was 12 ± 2% of radioligand added (1,700 Bq), and nonspecific binding was estimated to be 15–25% of the total [¹²⁵I]hCT binding. The binding data show that hCT is the least effective of all of the peptides tested in displacing [¹²⁵I]hCT, (Table 2 and Fig. 3a). The measured IC₅₀ for hCT is 3,220 ± 240 pM, whereas the affinity of the remaining variants (sCT, 6CHR, 6S, and 5P) is around an order of magnitude higher ($P < 0.01$), with the 5P variant having the highest affinity at 154 ± 34 pM. These data also indicate that the modifications introduced in the designs to prevent aggregation do not interfere with their binding to the CT receptor.

The stimulation of the CT receptor by the CT variants was quantified by monitoring the intracellular cAMP levels before and after incubation of T47D cells with each of the peptides. All of the peptides studied maximally stimulated cAMP accumulation between 300- and 360-fold over basal levels (2–6 pmol per 500,000

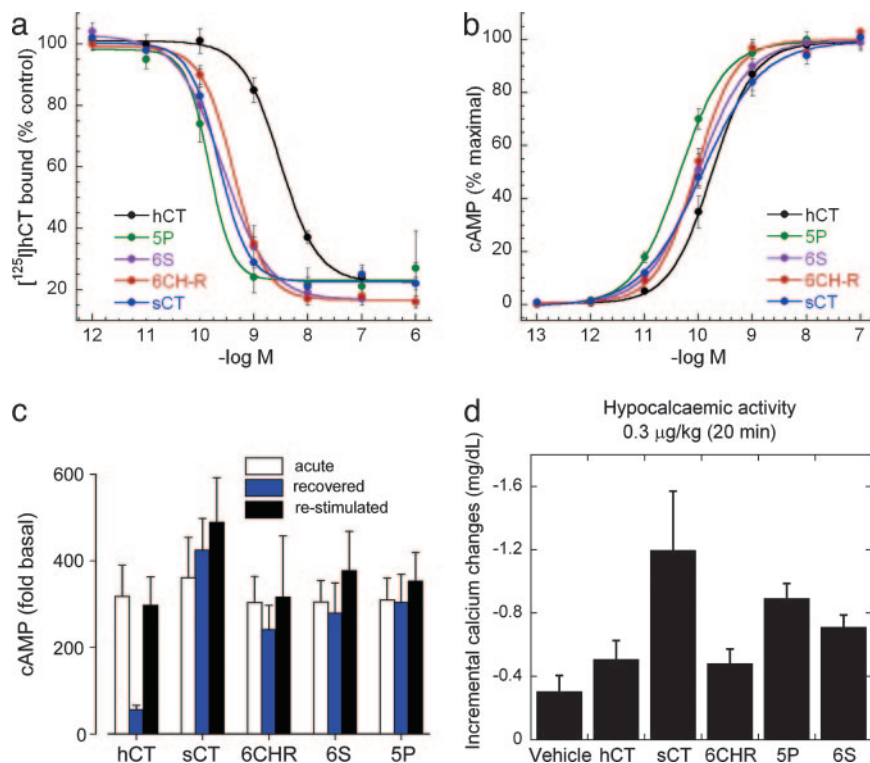


Fig. 3. Receptor binding and stimulation activity of CT variants. (a) [¹²⁵I]hCT binding inhibition measured in T47D cells. Cells were incubated for 5 h at 15°C with 1,700 Bq [¹²⁵I]hCT in the absence (control vehicle) or presence of different CT variants. Total binding was 197 ± 27 Bq per 500,000 cells and nonspecific binding was 15–25% of total binding. Results are means ± SEM (standard error mean) from three independent experiments. (b) Stimulation of intracellular cAMP levels after incubation in the presence of CT variants or vehicle alone. Cells were incubated for 15 min at 37°C in the presence of 1 mM IBMX. Basal cAMP levels ranged from 2–6 pmol per 500,000 cells and maximal cAMP levels from 500–3,000 pmol per 500,000 cells. The results shown represent the means ± SEM corresponding to six independent experiments. (c) Persistent stimulation of T47D cells. Cells were pretreated with a saturating concentration of unlabelled peptides (100 nM) under similar conditions to those used for the measurement of cAMP stimulation, except that IBMX was omitted during all this initial period. Cells were then washed extensively for a total period of 6 h, and cAMP levels were then measured in the presence of IBMX after stimulation with 100 nM of each of the various peptides in T47D cells previously subjected to a pretreatment with the different peptides (re-stimulated), as well as in cells that had not been exposed to such pretreatment. An additional group of samples corresponded to pretreated T47D cells that had not been subjected to a second incubation with the different CT variants (recovered). Results are means ± SEM of four independent experiments. The observation of similar cAMP levels in acutely stimulated cells and restimulated cells with all of the peptide variants indicates that in all cases no significant receptor down-regulation has occurred during the chase period. In all cases, the control corresponds to cells incubated in the presence of vehicle only. (d) Hypocalcaemic activity of CT variants on rats. Values correspond to the variation in calcium levels in blood (means ± SEM) 20 min after the administration of each of the CT variants at 0.3 μg/kg and the vehicle alone.

cells). This assay also showed hCT to have the lowest potency of all of the CT variants, stimulating cAMP production with an EC_{50} of 231 ± 73 pM (Fig. 3*b* and Table 2). sCT, 6CH-R, and 6S were approximately equally potent, being on average twice as effective as hCT in this assay; whereas 5P was 2-fold more potent than sCT, 6CH-R, and 6S, and 5-fold more potent than hCT ($P < 0.05$). The increase in efficacy observed for the engineered variants is likely to arise as a direct consequence of their reduced propensity to aggregate and increased availability, as previous studies seem to suggest (24). This possibility was further investigated by measuring the persistence of cAMP stimulation in T47D in the presence of the different peptides. All of the engineered CT variants seem to persist in stimulating the accumulation of cAMP more effectively than hCT, once the peptides are withdrawn from the media (Fig. 3*c*). This finding can be related to longer-term receptor activation, as it is the case for sCT, perhaps associated with an increased facility of the CT variants to interact with membrane lipids (42, 43). Such an observation could have important repercussions in the use of these reengineered variants as therapeutics. Improved pharmacokinetics should enable the use of lower CT doses (reducing the risk of undesired side effects) or to decrease dosing frequency.

The activity of the different CT variants was further analyzed by monitoring their hypocalcaemic effect in a rat model (Fig. 3*d*). Despite the limitations presented by this model [existence of renal effects that stimulate calcium secretion in urine, rather than inhibition of bone resorption, and differences in binding to rat CT receptors compared to the specific hCT_(a) receptor], we observed that two of the redesigned CT variants (6S and 5P) have a stronger hypocalcaemic effect than hCT 20 min after their administration to rats. Moreover, 5P and sCT are shown to have comparable hypocalcaemic effects (within the large standard errors observed), in these tests.

Further characterization of the efficacy of these variants would require testing their biological activity in relevant animal models [e.g., transgenic animals expressing hCT_(a) receptors], bioavailability through different routes of administration (injection vs. nasal or even oral), pharmacokinetics, and immunogenicity (e.g., cross-reactivity of antibodies found in patients subjected to long term sCT therapy, or human T cell stimulation by the different peptides). All of these are beyond the scope of the present study, which is to prove that rational modification of polypeptides can be successfully used to create active variants that exhibit a substantially reduced aggregation propensity.

It is also very important to note that the present approach can be used to address protein aggregation, regardless of the type of assembly formed. Indeed, these algorithms have been tested in proteins under conditions in which amyloid fibrils were not formed (cases of oligomerization, protofibril formation, or early stages of amyloid formation in which amorphous aggregates are formed) (26), as well as with nonamyloidogenic proteins that only seem to produce amorphous assemblies (27) (A. Pawar, J.Z., F.C., S.B.F., S.P., C. Wright, and M. Vendruscolo, unpublished observations). Therefore, the aggregation propensity of polypeptides that self-associate but do not necessarily form amyloid fibrils can be predicted with equal efficacy using similar approaches. This finding shows the general applicability of this approach, as suggested by studies on several other polypeptides (31, 44), including those with therapeutic biological relevance (S.P., S.B.F., and J.Z., unpublished results). Further implementation and improvement of the algorithms used here, as well as those developed by other groups, will

be valuable for successful protein redesign aimed to prevent aggregation (26–28, 39, 45, 46).

Conclusions

We have successfully designed three hCT-like variants with increased stability in solution using a semiempirical approach based on previous studies on model amyloid proteins (30, 31, 39, 47), and a more quantitative approach involving an *in silico* selection procedure that makes use of a combination of physicochemical parameters to describe the aggregation propensities of polypeptide chains (26–28, 39). The second methodology increased dramatically the throughput of the screening process, enabling the selection of candidates from a large number of sequences involving the simultaneous mutation of multiple residues. The aggregation propensity of the engineered variants was found to be substantially decreased when compared to the human sequence, especially in those variants designed using the predictive algorithm (5P and 6S). These variants also clearly outperform sCT in the aggregation assays. Most importantly, all redesigned variants show increased physiological activity when compared to hCT, both in terms of receptor binding and cAMP stimulation. Although two (6CHR and 6S) show comparable activity to sCT, one of the variants (5P) shows 2- to 3-fold higher potency than the salmon variant, the currently favored therapeutic form of CT. Indeed two of the reengineered CT variants (5P and 6S) show an enhanced hypocalcaemic activity *in vivo* when compared with that of hCT, further validating our design approach.

The substantially reduced number of sequence differences from hCT in these reengineered variants compared to sCT, together with their lower-aggregation propensities, also should minimize the risk of undesired immunogenic responses (19–23) or gastric dysfunctions, such as anorexia, nausea, and vomiting (17), that have been reported to occur in patients under treatment with sCT. It has been suggested that the first nine residues at the N terminus of sCT could be particularly important in promoting anorexia (18). Because all of the redesigned hCT variants described in this study have complete sequence identity with the first 11 residues of the hCT sequence, they could well have an improved gastric tolerance on the basis of these previous studies (18).

The success of the present study on CT, together with previous evidence showing the general applicability of these algorithm to predict the aggregation behavior of a large collection of polypeptides (26, 27), suggests that the approach described here could facilitate the general redesign of polypeptide therapeutics to produce a much higher resistance to aggregation, and therefore a higher bioavailability, than natural sequences. Therefore, it could also be extremely useful in the development of more user-friendly formulations for many drugs that currently require parental administration. We believe, on the basis of the evidence from this study, that the present methodology could have a significant impact as a general strategy for designing new forms of polypeptide-based biopharmaceuticals. The major aims of such an approach would be to optimize production, formulation, posology, and shelf-life for these polypeptide drugs, and perhaps even to eliminate the need of cold-delivery chains. All these could indeed have important repercussions in both the therapeutic efficacy of such molecules and the overall economics of drug discovery.

We thank K. Husmann for the preparation of iodinated human calcitonin and P. Ledger for suggestions to produce Fig. 1*b*. This work was funded through a Research Grant from Zyentia, Ltd. J.Z. also acknowledges support from the Leverhulme Trust.

1. Caldwell, G. W., Ritchie, D. M., Masucci, J. A., Hageman, W. & Yan, Z. (2001) *Curr. Top. Med. Chem.* **1**, 353–366.
2. Chi, E. Y., Krishnan, S., Randolph, T. W. & Carpenter, J. F. (2003) *Pharm. Res.* **20**, 1325–1336.
3. Cleland, J. L., Powell, M. F. & Shire, S. J. (1993) *Crit. Rev. Ther. Drug Carrier Syst.* **10**, 307–377.

4. Schellekens, H. (2002) *Nat. Rev. Drug Discov.* **1**, 457–462.
5. Silverman, S. L. (1997) *Am. J. Med. Sci.* **313**, 13–16.
6. Schneider, D., Hofmann, M. T. & Peterson, J. A. (2002) *Am. Fam. Physician* **65**, 2069–2072.
7. Mehta, N. M., Malootian, A. & Gilligan, J. P. (2003) *Curr. Pharm. Des.* **9**, 2659–2676.

8. Sletten, K., Westermark, P. & Natvig, J. B. (1976) *J. Exp. Med.* **143**, 993–998.
9. Silver, M. M., Hearn, S. A., Lines, L. D. & Troster, M. (1988) *J. Histochem. Cytochem.* **36**, 1031–1036.
10. Arvinte, T., Cudd, A. & Drake, A. F. (1993) *J. Biol. Chem.* **268**, 6415–6422.
11. Zaidi, M., Inzerillo, A. M., Moonga, B. S., Bevis, P. J. & Huang, C. L. (2002) *Bone* **30**, 655–663.
12. Schubert, D., Behl, C., Lesley, R., Brack, A., Dargusch, R., Sagara, Y. & Kimura, H. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 1989–1993.
13. Braun, A., Kwee, L., Labow, M. A. & Alsenz, J. (1997) *Pharmacol. Res.* **14**, 1472–1478.
14. Brange, J., Andersen, L., Laursen, E. D., Meyn, G. & Rasmussen, E. (1997) *J. Pharmacol. Sci.* **86**, 517–525.
15. Rymer, D. L. & Good, T. A. (2000) *J. Biol. Chem.* **276**, 2523–2530.
16. Curatolo, L., Valsasina, B., Caccia, C., Raimondi, G. L., Orsini, G. & Bianchetti, A. (1997) *Cytokine* **9**, 734–739.
17. Feletti, C. & Bonomini, V. (1979) *Nephron* **24**, 85–88.
18. Yamamoto, Y., Nakamura, H., Koida, M., Seyler, J. K. & Orłowski, R. C. (1982) *Jpn. J. Pharmacol.* **32**, 1013–1017.
19. Levy, F., Muff, R., Dotti-Sigris, S., Dambacher, M. A. & Fischer, J. A. (1988) *J. Clin. Endocrinol. Metab.* **67**, 541–545.
20. Muff, R., Dambacher, M. A. & Fischer, J. A. (1991) *Osteoporosis Int.* **1**, 72–75.
21. Grauer, A., Ziegler, R. & Raue, F. (1995) *Exp. Clin. Endocrinol. Diabetes* **103**, 345–351.
22. Rodriguez, A., Trujillo, M. J., Herrero, T., Baeza, M. L. & de Barrio, M. (2001) *Allergy* **56**, 801.
23. Porcel, S. L., Cumplido, J. A., de la Hoz, B., Cuevas, M. & Losada, E. (2000) *Allergol. Immunopathol.* **28**, 243–245.
24. Cudd, A., Arvinte, T., Das, R. E., Chinni, C. & MacIntyre, I. (1995) *J. Pharmacol. Sci.* **84**, 717–719.
25. Chiti, F., Calamai, M., Taddei, N., Stefani, M., Ramponi, G. & Dobson, C. M. (2002) *Proc. Natl. Acad. Sci. USA* **99**, Suppl. 4, 16419–16426.
26. Chiti, F., Stefani, M., Taddei, N., Ramponi, G. & Dobson, C. M. (2003) *Nature* **424**, 805–808.
27. DuBay, K. F., Pawar, A. P., Chiti, F., Zurdo, J., Dobson, C. M. & Vendruscolo, M. (2004) *J. Mol. Biol.* **341**, 1317–1326.
28. Pawar, A., Dubay, K. F., Zurdo, J., Chiti, F., Dobson, C. M. & Vendruscolo, M. (2005) *J. Mol. Biol.* **350**, 379–392.
29. Zurdo, J., Guijarro, J. I., Jimenez, J. L., Saibil, H. R. & Dobson, C. M. (2001) *J. Mol. Biol.* **311**, 325–340.
30. Villegas, V., Zurdo, J., Filimonov, V. V., Aviles, F. X., Dobson, C. M. & Serrano, L. (2000) *Protein Sci.* **9**, 1700–1708.
31. Ventura, S., Zurdo, J., Narayanan, S., Parreno, M., Mangués, R., Reif, B., Chiti, F., Giannoni, E., Dobson, C. M., Aviles, F. X. & Serrano, L. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 7258–7263.
32. Zimmermann, U., Fluchmann, B., Born, W., Fischer, J. A. & Muff, R. (1997) *J. Endocrinol.* **155**, 423–431.
33. Hunter, W. M. & Greenwood, F. C. (1962) *Nature* **194**, 495–496.
34. Hussain, A. A., Jona, J. A., Yamada, A. & Dittert, L. W. (1995) *Anal. Biochem.* **224**, 221–226.
35. Reches, M., Porat, Y. & Gazit, E. (2002) *J. Biol. Chem.* **277**, 35475–35480.
36. Kazantzis, A., Waldner, M., Taylor, J. W. & Kapurniotu, A. (2002) *Eur. J. Biochem.* **269**, 780–791.
37. Conner, A. C., Hay, D. L., Howitt, S. G., Kilk, K., Langel, U., Wheatley, M., Smith, D. M. & Poyner, D. R. (2002) *Biochem. Soc. Trans.* **30**, 451–455.
38. Kallberg, Y., Gustafsson, M., Persson, B., Thyberg, J. & Johansson, J. (2001) *J. Biol. Chem.* **276**, 12945–12950.
39. Zurdo, J. (2005) *Protein Pept. Lett.* **12**, 171–187.
40. Beaudreuil, J., Taboulet, J., Orcel, P., Graulet, A. M., Denne, M. A., Baudoin, C., Jullienne, A. & De Vernejoul, M. C. (2000) *Bone* **27**, 161–168.
41. Poyner, D. R., Sexton, P. M., Marshall, I., Smith, D. M., Quirion, R., Born, W., Muff, R., Fischer, J. A. & Foord, S. M. (2002) *Pharmacol. Rev.* **54**, 233–246.
42. Schmidt, M. C., Rothen-Rutishauser, B., Rist, B., Beck-Sickinger, A., Wunderli-Allenspach, H., Rubas, W., Sadee, W. & Merkle, H. P. (1998) *Biochemistry* **37**, 16582–16590.
43. Stipani, V., Callucci, E., Micelli, S., Picciarelli, V. & Benz, B. (2001) *Biophys. J.* **81**, 3332–3338.
44. Frare, E., Polverino De Laureto, P., Zurdo, J., Dobson, C. M. & Fontana, A. (2004) *J. Mol. Biol.* **340**, 1153–1165.
45. Fernandez-Escamilla, A. M., Rousseau, F., Schymkowitz, J. & Serrano, L. (2004) *Nat. Biotechnol.* **22**, 1302–1306.
46. Tartaglia, G. G., Cavalli, A., Pellarin, R. & Caflich, A. (2004) *Protein Sci.* **13**, 1939–1941.
47. Lopez De La Paz, M., Goldie, K., Zurdo, J., Lacroix, E., Dobson, C. M., Hoenger, A. & Serrano, L. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 16052–16057.
48. Munoz, V. & Serrano, L. (1994) *Nat. Struct. Biol.* **1**, 399–409.
49. Broome, B. M. & Hecht, M. H. (2000) *J. Mol. Biol.* **296**, 961–968.