Purpose of Supplement

This Supplement contains additional NMR data, structural comparisons and associated calculations in support. Nine supplemental figures (supplemental Fig. S1-S9) and three supplemental tables (Tables S1-S3) are provided.

Table S3 gives methyl proton secondary shifts for Val, Leu, and Ala residues of SCI-b and a two-chain control analog [Asp^{B10}, Lys^{B28}, Pro^{B29}]-insulin (DKP-insulin; see supplemental ref. (1)). These tables are motivated by the sensitivity of ^{1}H methyl resonances to neighboring aromatic-ring current shifts and hence tertiary structure. The methyl resonances of Leu^{B15}, for example, are shifted upfield due to the aromatic ring current of Phe^{B24}. Ring-current-shift calculations pertaining to insulin were reported by Jacoby et al. in supplemental ref. (2).

Supplemental Figure S1 presents ^{1}H-^{15}N- and ^{1}H-^{13}C-HSQC spectra as well as selected planes of a 4D time-shared NOESY spectrum. The latter provided inter-residue NOE cross-peaks (or lack thereof) as supporting evidence for structural features identified in the SCI-b simulated-annealing ensemble (Fig. 2 in the main text). Figure S2 illustrates the side-chain environments of engineered sites in the SCI-b ensemble (His^{A8}, Glu^{A14}, Asp^{B28} and Pro^{B29}). Figure S3 presents DLS studies of SCI-b and insulin *lispro* under formulation conditions at 0.6 mM (*i.e.*, nominal strength U-100 in international units) before and after heating samples at 88 °C for 10 min. These data demonstrate that, whereas SCI-b was unchanged by such thermal stress, insulin *lispro* formed large aggregates, thereby rationalizing the impaired refolding observed in CD-monitored temperature denaturation studies (Fig. 5*C* in the main text). Supplemental Figure S4 contains strip plots from the 3D HNCO triple resonance spectrum of [^{13}C, ^{15}N]-labeled SCI-b whose pattern of cross-peak intensities demonstrate the presence of motional narrowing in residues B30, C1, and C2 (relative to core residues B13, B14, and B18). In Figure S5 we compare dynamical information derived from chemical-shift analyses of SCI-b versus DKP-insulin (as a representative two-chain analog). This comparison provides evidence of (a) damping of conformational fluctuations in the B25-B30 segment due to the presence of the C domain in SCI-b and (b) improved secondary structure propensity (SSP) scores of the A1-A8 helix by Thr^{A8}→His substitution.

The main text presents alternative ensembles of SCI-a and SCI-b based on distinct assumptions. These are intended to address a thought ("*gedanken*") experiment: what would a nanoscale Maxwell's Demon see in a physical ensemble? Figure S6 illustrates the procedure for our multi-conformer simulations, which utilize X-ray diffraction data of related analog SCI-a (see companion study (3)) to generate ensemble models of the SCI-a hexamer. In Supplemental Figure S7 are given average per-residue RMSDs of these alternative ensembles (Fig. 8 in the main text).

IGF-I and IGF-II provide examples of "SCIs" whose evolution reflects gene duplication and specialization in vertebrates. In Figure S8 is shown an overlay of IGF-I and IGF-II NMR ensembles with SCI-b (ensemble from Fig. 2 in the main text), showing similar segmental imprecision in their respective C domains.

Figure S9 illustrates insulin's biphasic secretion from β-cells in support of our Discussion, highlighting the rationale and potential utility of ultra-stable biphasic insulin therapeutics.

Supplemental Table S1. Statistics of SCI-b ensemble (30 best-fit structures)

Parameter	SA Ensemble[a]	Minimized Average
RMSD (Å) from experimental restraints (by category)[b]		
All (1370)	0.05 ± 0.01	0.04
Intraresidue, i=j (633)	0.039 ± 0.001	0.04
Sequential, |i-j|=1 (285)	0.048 ± 0.004	0.04
Medium range, 1<|i-j|<5 (235)	0.044 ± 0.002	0.04
Long range, |i-j|≥5 (217)	0.05 ± 0.04	0.02
RMSD from idealized covalent geometry		
Bonds (Å)	0.006 ± 0.001	0.005
Angles (°)	0.64 ± 0.04	0.56
Impropers (°)	0.49 ± 0.01	0.39
Energy (kcal/mol)[c]		
	-178.0 ± 8.4	-
Ramachandran plot dihedral angle distributions[d]		
Most favored regions (%)	82.5	82.5
Additionally and generously allowed regions (%)	17.5	17.5
Disallowed regions (%)	0.0	0.0
Coordinate precision of insulin core region[e]		
RMSD of backbone atoms to mean (Å)	0.39	
RMSD of all heavy atoms (sidechain) to mean (Å)	0.79	
Coordinate precision of B27-C4 region[f]		
RMSD of backbone atoms to mean (Å)	1.14	
RMSD of all heavy atoms (sidechain) to mean (Å)	1.60	

[a]Mean ± standard error

[b]Numbers in parentheses report number of distance restraints in each category. RMSD values were calculated using all residues in the 30 models.

[c]Lennard-Jones potential energy as calculated with CHARMM19 using empirical energy parameters

[d]Using PROCHECK-NMR (4). Only regions B5-B26, C4-C6, and A1-A20 (for all models in the ensemble) were considered when calculating these statistics as all other residues resided in flexible regions of the ensemble. Glycine residues not included.

[e]Residues B5-B26,C4-C6,A1-A20

[f]Residues B27-C4

Supplemental Table S2. ^{1}H-^{2}H Exchange Studies of [^{15}N]-SCI-b (pH 2.8 and 25 ^{o}C)

Residue	k_{obs} ($s^{-1} \times 10^{6}$)	k_{int} ($s^{-1} \times 10^{6}$)	PF	$\Delta G_{pH2.8}$ (kcal/mol)
B6	220.5 ± 9.9	1230	5.6 ± 0.2	1.02 ± 0.03
B11	38.7 ± 1.8	1230	31.8 ± 1.5	2.05 ± 0.03
B12	9.3 ± 0.2	1000	107.2 ± 2.6	2.77 ± 0.01
B14	2.7 ± 0.3	2080	760.7 ± 77.0	3.93 ± 0.06
B18[a]	0.3 ± 0.2	1000	3182.0 ± 1580.0	4.77 ± 0.29
B19	1.7 ± 0.2	1380	797.4 ± 97.8	3.95 ± 0.07
B23	369.7 ± 94.7	1790	4.8 ± 1.2	0.93 ± 0.15
A10	79.2 ± 24.4	914	11.5 ± 3.6	1.45 ± 0.18
A12	46.0 ± 0.6	2160	46.9 ± 0.6	2.28 ± 0.01
A13	294.0 ± 9.4	1040	3.5 ± 0.1	0.75 ± 0.02
A14	81.2 ± 1.0	1280	15.8 ± 0.2	1.63 ± 0.01
A15	3.6 ± 0.2	1230	337.3 ± 18.2	3.44 ± 0.03
A16[a]	0.2 ± 0.1	1070	4463.0 ± 1013.3	4.97 ± 0.13
A17	2.0 ± 0.1	1240	606.5 ± 16.6	3.79 ± 0.02
A19[a]	0.5 ± 0.1	1430	3015.9 ± 564.9	4.74 ± 0.11
A20	4.0 ± 0.2	1620	408.2 ± 19.1	3.56 ± 0.03

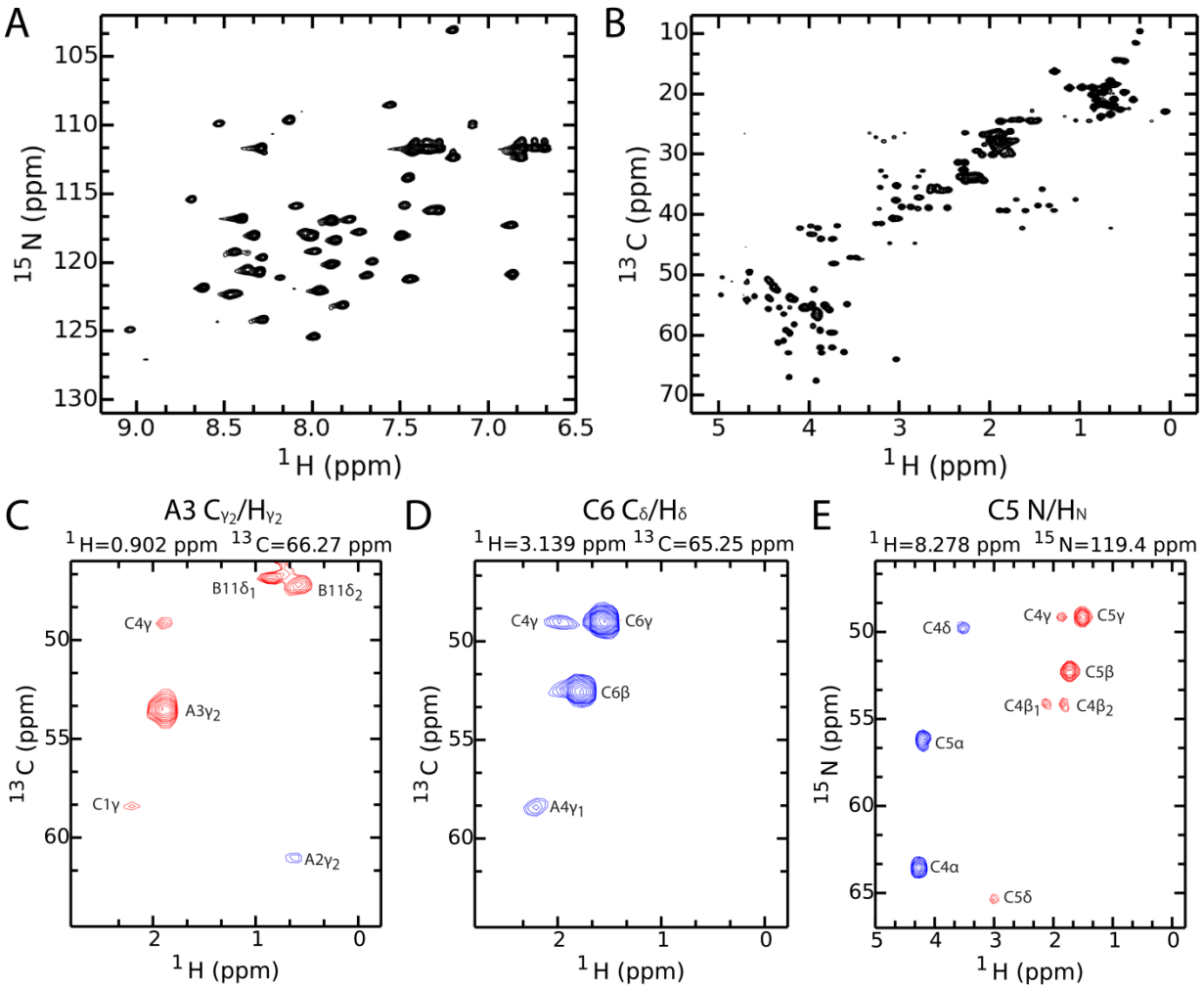
[a]The exponential fit for B18 resonance was spurious due to weak starting peak intensity and very slow exchange. Global resonances A16 and A19 also decay very slowly, resulting in relatively large error. Though all three globals give approximately the correct $\Delta G_{pH2.8}$ that agrees with CD data, this experiment is intended to probe subglobal exchange only.

Supplemental Table S3. ^{1}H secondary shifts[a] for SCI-b and DKP-insulin[b] control

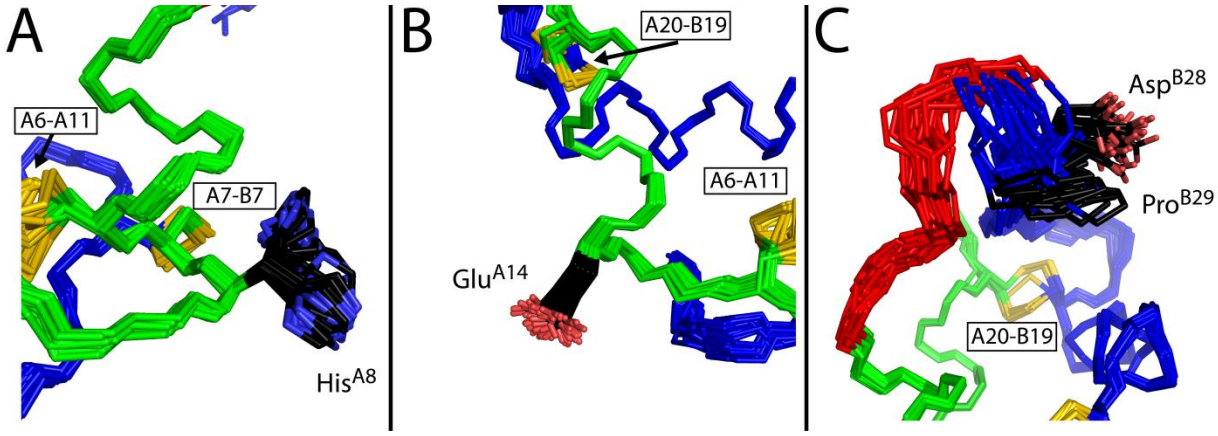
B chain Residue	Atom Name	SCI-b	DKP
Val^{B2}	Hβ	-0.14	-0.21
Val^{B2}	Hγ1	-0.09	-0.20
Val^{B2}	Hγ2	-0.09	-0.14
Leu^{B6}	Hβ1	0.20	-0.93
Leu^{B6}	Hβ2	-0.78	0.08
Leu^{B6}	Hδ1	-0.11	-0.11
Leu^{B6}	Hδ2	0.08	-0.18
Leu^{B6}	Hγ	0.13	0.01
Leu^{B11}	Hβ1	0.27	0.15
Leu^{B11}	Hβ2	-0.39	-0.50
Leu^{B11}	Hδ1	-0.23	-0.33
Leu^{B11}	Hδ2	-0.13	-0.23
Leu^{B11}	Hγ	-0.36	-0.42
Val^{B12}	Hβ	0.04	0.00
Val^{B12}	Hγ1	0.11	-0.01
Val^{B12}	Hγ2	0.00	0.04
Ala^{B14}	Hβ	0.07	-0.07
Leu^{B15}	Hβ1	-0.63	-1.54
Leu^{B15}	Hβ2	-1.45	-0.74
Leu^{B15}	Hδ1	-0.67	-0.82
Leu^{B15}	Hδ2	-0.28	-0.39
Leu^{B15}	Hγ	-0.24	-0.38
Leu^{B17}	Hβ1	0.38	0.27
Leu^{B17}	Hβ2	0.17	0.11
Leu^{B17}	Hδ1	0.06	0.00
Leu^{B17}	Hδ2	0.16	0.03
Leu^{B17}	Hγ	0.27	0.15
Val^{B18}	Hβ	-0.07	-0.17
Val^{B18}	Hγ1	0.10	-0.02
Val^{B18}	Hγ2	-0.02	0.04

A chain Residue	Atom Name	SCI-b	DKP
Ile^{A2}	Hβ	-0.73	-0.80
Ile^{A2}	Hδ	-0.29	-0.42
Ile^{A2}	Hγ1,1	-0.29	-0.41
Ile^{A2}	Hγ1,2	-0.29	-0.36
Ile^{A2}	Hγ2	-0.13	-0.30
Val^{A3}	Hβ	-0.02	-0.22
Val^{A3}	Hγ1	-0.12	-0.13
Val^{A3}	Hγ2	0.08	-0.05
Ile^{A10}	Hβ	-0.27	-0.35
Ile^{A10}	Hδ	-0.35	-0.43
Ile^{A10}	Hγ1,1	-0.37	-0.42
Ile^{A10}	Hγ1,2	-0.77	-0.82
Ile^{A10}	Hγ2	-0.23	-0.31
Leu^{A13}	Hβ1	0.01	-0.25
Leu^{A13}	Hβ2	-0.08	-0.19
Leu^{A13}	Hδ1	-0.01	-0.11
Leu^{A13}	Hδ2	0.02	-0.12
Leu^{A13}	Hγ	0.06	-0.11
Leu^{A16}	Hβ1	-0.15	0.31
Leu^{A16}	Hβ2	0.45	-0.28
Leu^{A16}	Hδ1	-0.07	-0.25
Leu^{A16}	Hδ2	-0.05	-0.17
Leu^{A16}	Hγ	0.26	0.10

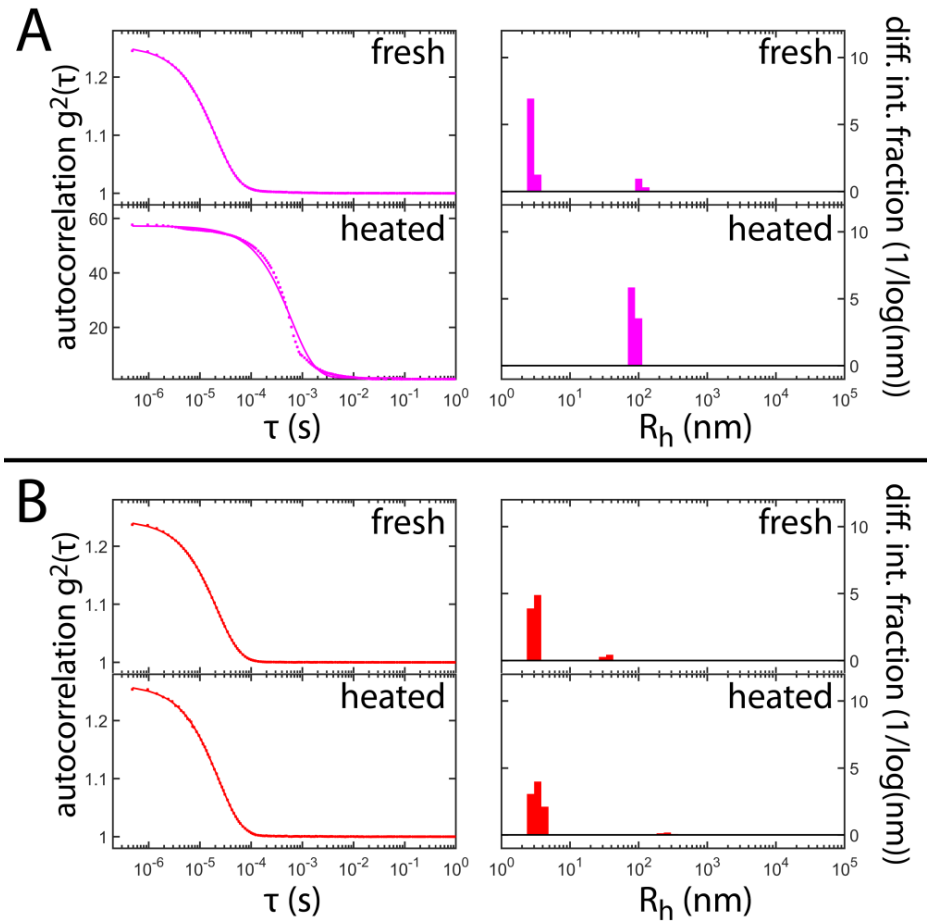
[a]Defined as methyl proton chemical shift minus the tabulated random coil shifts in supplemental ref. (5).
[b]DKP refers to two chain [Asp^{B10},Lys^{B28},Pro^{B29}]-insulin, whose chemical shifts were derived from prior studies (supplemental ref. (1))



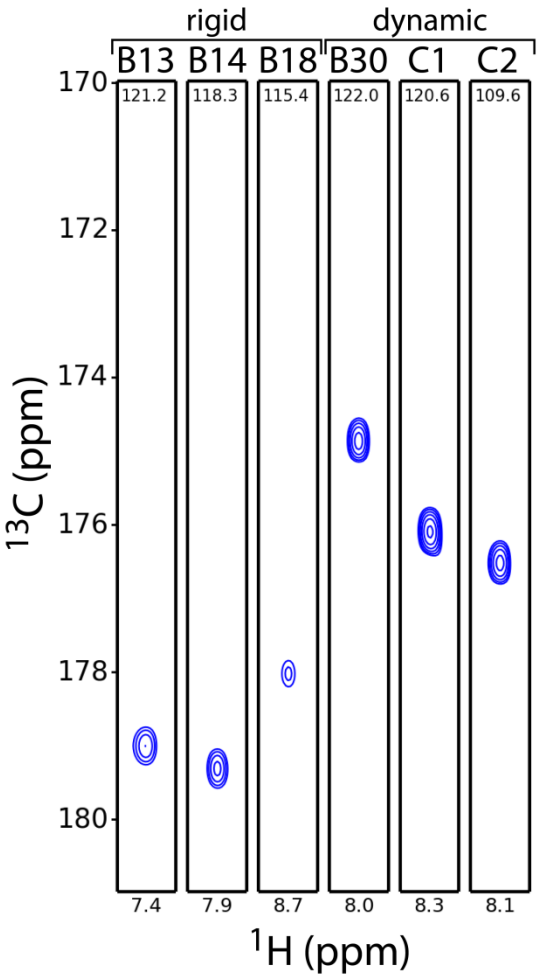
Supplemental Figure S1. NMR spectra of [^{13}C,^{15}N]-SCI-b at pH 7.4 and 25°C. (*A*) ^{1}H-^{15}N- and (*B*) ^{1}H-^{13}C- HSQCs of [^{13}C, ^{15}N]-labeled SCI-b. Panels (*C-D*) and panel (*E*) are planes from the [^{13}C, ^{13}C]-4D-NOESY and [^{15}N, ^{13}C]-4D-NOESY (respectively) for the resonances stated at top of each plot. NOE cross-peaks between C4γ and A3γ protons in (*C*) suggest a compact C domain that brings Pro^{C4} within NOE range of Val^{A3}. Appearance of the A4γ_1 resonance in (*D*) reflects the presence of the Glu^{A4}-Arg^{C6} salt bridge. Panel (*E*) highlights the lack of any medium-range NOEs in the C5 amide plane. Spectra were acquired at pH 7.4 and 25 °C at a protein concentration of ~0.6 mM.



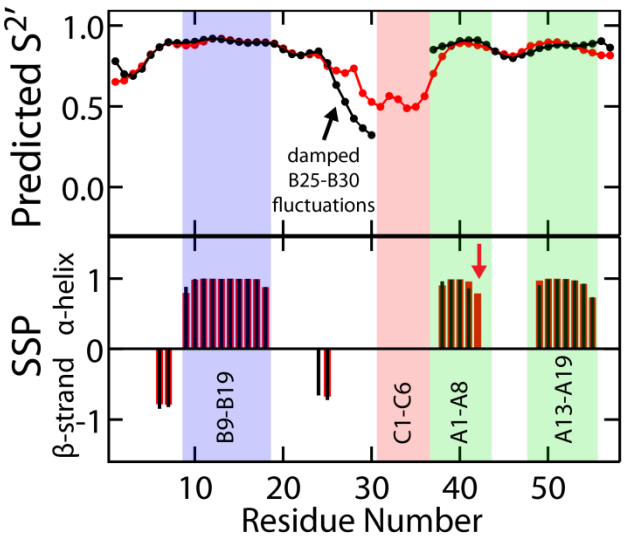
Supplemental Figure S2. Side-chain environments of B28, B29, A8 and A14. An ensemble (30 structures) of SCI-b derived from Figure 2 (main text) focusing on (*A*) His^{A8}, (*B*) Glu^{A14}, and (*C*) Asp^{B28} and Pro^{B29}. Disulfides are labeled in *boxes*. Color code: A domain, *green*; B domain, *blue*; C domain, *red*; disulfides, *yellow*; oxygen, *light red*; nitrogen, *light blue;* and side-chain carbons, *black*.



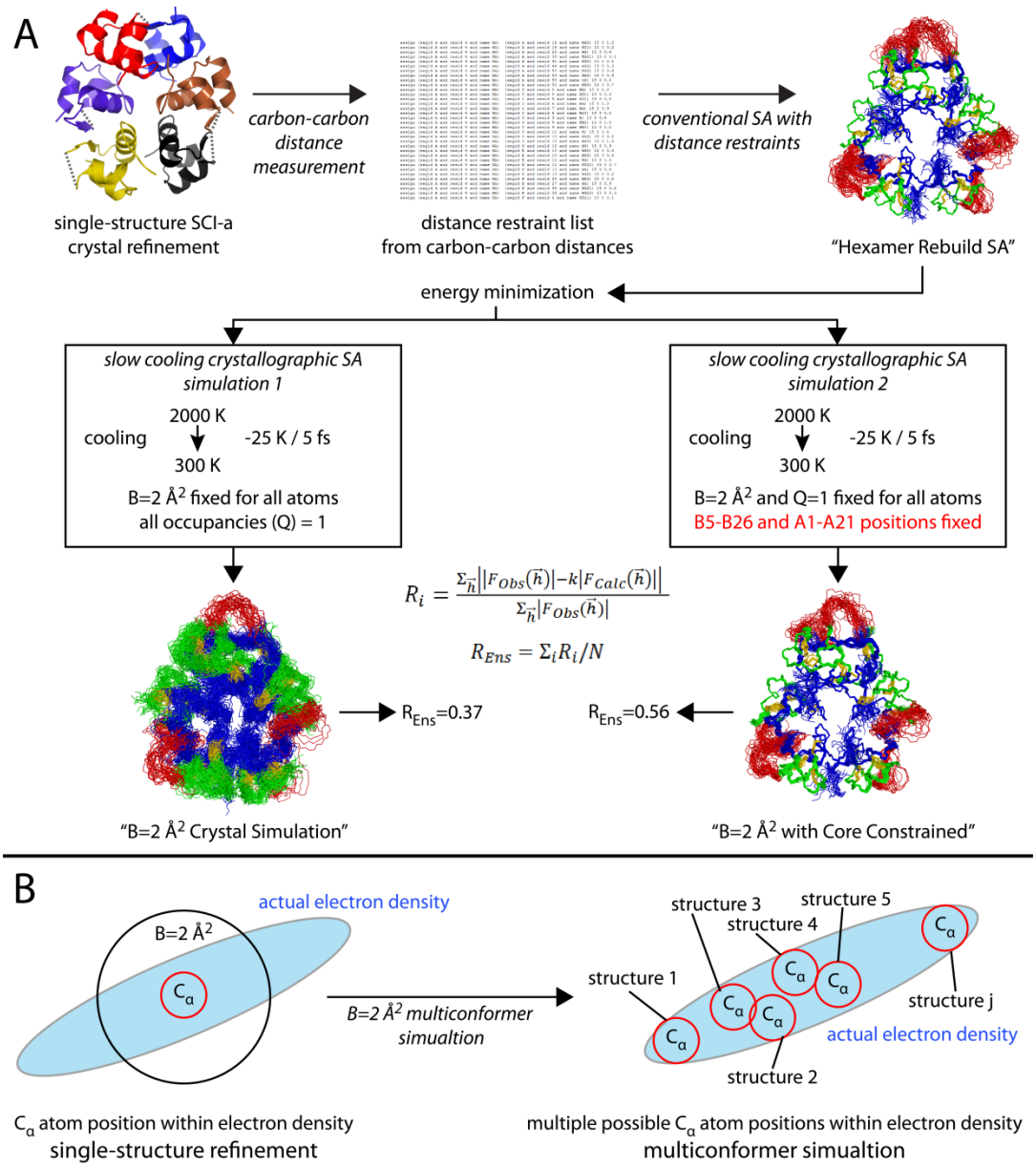
Supplemental Figure S3. Heat-induced aggregation of *lispro* and SCI-b probed by DLS. Dynamic light scattering of (*A*) insulin *lispro* (*magenta*) and (*B*) SCI-b (*red*) at U-100 strength (0.6 mM) in Tris-HCl buffer (pH 7.4 at room temperature) supplemented with 0.3 mM ZnCl. DLS samples were either freshly prepared or evaluated following incubation at 88 ^{o}C for 10 min. Autocorrelation functions (*left panel*) and differential intensity fraction-size distributions (*right panel*) disclose the formation of very large aggregates in heated *lispro* samples whereas SCI-b is unchanged by incubation at 88 °C. Fresh *lispro* and SCI-b each give the predominance of hydrodynamic radii (R_h) characteristic of lower molecular-mass species (insulin monomers, dimers and hexamers), in accordance with our companion study (3). Aggregates formed in the heated *lispro* sample give rise to a slowly decaying autocorrelation function with high A-to-B ratio (60 to 1, whereas the theoretical maximum within the Rayleigh scattering limit is 2 to 1), reflecting the slow diffusion of large, intensely scattering species into and out of the laser beam. These data suggest heat-induced aggregation of *lispro* is responsible for its irreversible unfolding as observed in CD-monitored temperature scans (Fig. 5*C* of main text).



Supplemental Figure S4. Motional narrowing within C domain of SCI-b. Strip plots from 3D HNCO spectrum of [^{13}C,^{15}N]-SCI-b that exhibit greater carbonyl carbon-peak intensities of selected residues (Thr^{B30}, Glu^{C1} and Glu^{C2}) relative to core residues Glu^{B13}, Ala^{B14}, and Val^{B18}. *Small numbers* at top provide the chemical shift coordinate (in ppm) of the ^{15}N-dimension for each strip. This spectrum was acquired at pH 7.4 and 25 °C at a protein concentration of ~0.6 mM.

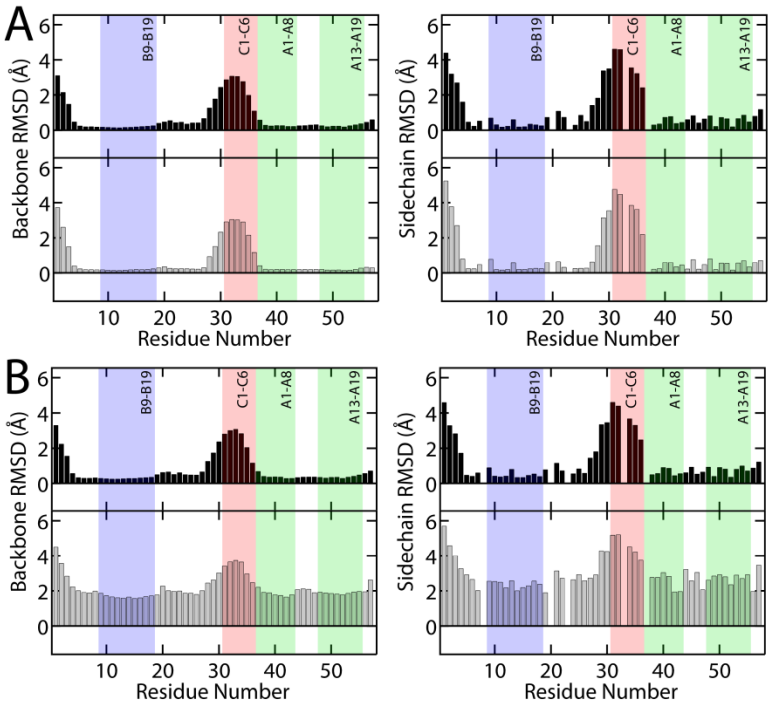


Supplemental Figure S5. Dynamic parameters of SCI-b *versus* DKP-insulin as a two-chain control. Predicted $S^{2'}$ (*top*) and SSP scores (*bottom*) for SCI-b (*red*) and DKP-insulin (*black*). Due to its absent C domain, DKP residues were numbered to match their counterpart in SCI-b (e.g., Gly^{A1} is residue 37 in both data sets) in each panel. Chemical shifts for DKP-insulin were obtained from prior studies (1); SSP scores and predicted $S^{2'}$ were generated by TALOS+ as in studies of SCI-b (see Methods). Damping of conformational fluctuations in the B25-B30 segment of SCI-b (indicated by higher predicted $S^{2'}$ values) are identified by the *black arrow* at *top*. Increased SSP helix-associated scores at the C-terminal end of the A1-A8 helix are highlighted by a *red arrow* in the *bottom panel*. DKP SSP scores are presented as *thin black bars* for visualization of overlapping SCI-b scores (*red bars*).

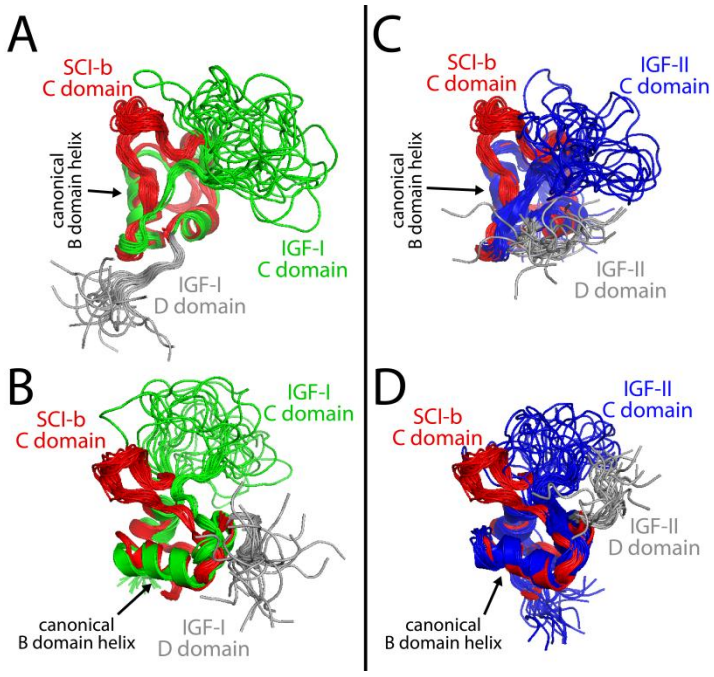


Supplemental Figure S6. Generation of multi-conformer SCI-a ensembles. (*A*) Flow chart in schematic form outlining protocol for SCI-a multi-conformer simulation and R_{Ens} calculation. First, the conventional single-structure SCI-a refinement (*top left corner*; see companion study(3)) was processed using a custom MATLAB script to measure all carbon-carbon distances between 1.5 Å–30 Å. From these distances, an XPLOR-formatted distance-restraint list (*middle*) was generated and used to create a 60-structure ensemble (labeled "Hexamer Rebuild SA") using the same simulated annealing (SA) protocol as for standard NMR structure determination. Each individual structure from this ensemble was energy-minimized and then subjected to two independent multi-conformer simulations. In *simulation 1* B-factors of all atoms were fixed to 2 Å^2 and all occupancies were set to 1 during a slow-cooling crystallographic SA (2000 K → 300 K with 25 K reduction per 5 fs time step). Here, B-factors of 2 Å^2 were chosen so that carbon atoms' mean square displacement (*u*, obtained from the equation $B=8\pi^2u^2$) would be just under 0.17 Å or ~10% the van der Waals radius of carbon (1.7 Å). *Simulation 2* was identical to simulation 1, but with the positions of residues B25-B26 and A1-A21,

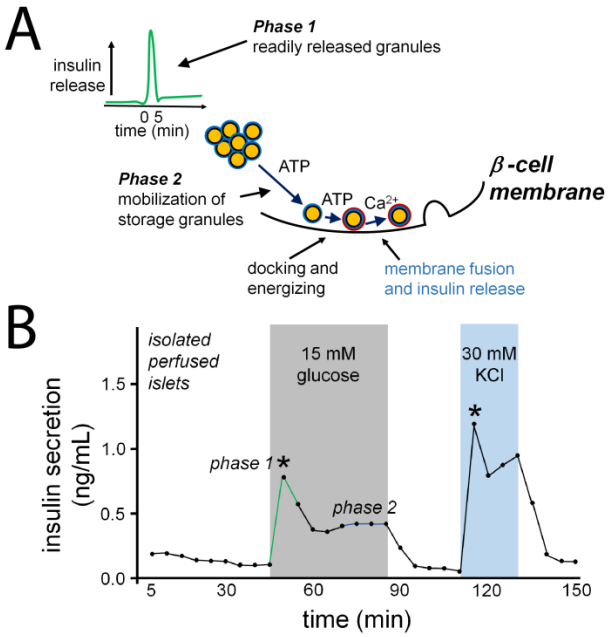
which gave well-defined electron density in the single-structure refinement, fixed during the simulation. Simulations *1* and *2* generated separate 60 structure ensembles labeled "B=2 Å^2 Crystal Simulation" and "B=2 Å^2 with Core Constrained." Ensemble-averaged R-factors, R_{Ens}, were calculated using the equation shown; the resulting values are given for each ensemble. R_{Ens} is higher for the "B=2 Å^2 with Core Constrained" ensemble (R_{Ens}=0.56) than that of the unrestrained "B=2 Å^2 Crystal Simulation" ensemble (R_{Ens}=0.37). Both ensemble R-factors are larger than that of the single-structure refinement (R_{Free}=0.31), likely due to the presence of C domain residues in the ensembles that have no corresponding electron density. The attenuated R_{Ens} of the more imprecise ensemble is a consequence of sampled atomic positions more completely fitting moderate-resolution electron density, a concept illustrated by panel (*B*). In panel (*B*) at *left*, the real position of a Cα atom (*red circle*) is circumscribed by a larger area defined by B=2 Å^2 (*black circle*). In a conventional single-structure crystal refinement, the atomic position fits within the actual electron density (represented as a non-spherical *blue* "smear"). By contrast, the B=2 Å^2 multi-conformer simulations generated a collection of possible C_α atom positions (*right*, one position per structure 1, 2, 3,…j) that also fit within the electron density, resulting in lower R_{Ens} in more imprecise ensembles. In the case of SCI-a, the R_{Ens} was not lower than the R_{Free} of the single-structure refinement due to addition of C domain residues with no corresponding electron density. The multi-conformer simulations, though "over-parameterizing" structural determination, provide "working hypotheses" as models of what the hexamer might look like to an observer existing on the nanoscale. The "B=2 Å^2 with Core Constrained" ensemble is perhaps the most physically realistic depiction of the hexameric ensemble as it reflects the rigidity of residues corresponding to well-defined electron density in conventional single-structure SCI-a refinement.



Supplemental Figure S7. RMSDs of alternative ensembles. Backbone (*left*) and side chains (*right*) RMSDs per residue: SCI-b (*black*) or SCI-a (*gray*) in the alternative ensembles shown in Figure 9 (main text). (*A*) SCI-b Original SA and SCI-a Hexamer Rebuild SA and (*B*) SCI-b All TA and SCI-a B=2 Crystal Simulation. RMSDs for the SCI-a hexamer ensembles are averaged over all monomers. Systematically high RMSD (≥2 Å) for the B=2 Crystal Simulation reflects a higher level of imprecision in this ensemble relative to counterpart time-averaged ensembles of SCI-b. The B9-B19, A1-A8, and A13-A19 helices and C1-C6 segment are highlighted by *shaded boxes*. RMSD calculations employed heavy atoms only. Glycines at position B8, B20, B23, C3, and A1 were excluded from side-chain RMSD calculations.



Supplemental Figure S8. Comparison of SCI-b with IGF-I and IGF-II. Structural alignments employed published NMR-derived structures of the IGFs. Superpositions of SCI-b ensemble (*red*; shown in Fig. 2 of the main text) to (*A-B*) IGF-I (*green*; PDB: 1PMX (6)) and (*C-D*) IGF-II (*blue*; PDB: 1IGL (7)). The canonical B9-B19 helix is identified (*black arrows*) to define the overall orientation of these proteins. Flexible IGF-I and IGF-II D domains are shown in *gray* to distinguish them from the flexible C domains. Alignments were performed according to the main-chain atoms of the three canonical helices (A1-A8, A13-A19, and B9-B19).



Supplemental Figure S9. **Biphasic β-cell insulin secretion and rationale for premixed soluble/microcrystalline insulin formulations**. *(A)* Phase 1 and phase 2 insulin secretion. Post-prandial rise in blood glucose is followed by biphasic insulin release: (*upper left*) Phase 1 is a rapid response requiring ATP and Ca^{2+} in which the "readily released granules" are secreted (within minutes); (*bottom right*) in delayed phase 2, storage granules (*orange* with *blue borders*) are moved from a granule storage reservoir, mobilized through docking (*orange* with *red borders*) and released. *(B)* Representative past studies of isolated perfused islets demonstrating phase-1 and phase-2 insulin secretion stimulated by glucose (*gray-shaded area*) or KCl (*blue*-shaded area) (adapted from ref. (8)). Asterisks highlight the phase 1 insulin secretion. Panel *(A)* is adapted from http://watcut.uwaterloo.ca/webnotes/Metabolism/Diabetes.html.

Supplemental References

1. Yang, Y., Petkova, A., Huang, K., Xu, B., Hua, Q.-x., Ye, I.-J., Chu, Y.-C., Hu, S.-Q., Phillips, N. B., Whittaker, J., Ismail-Beigi, F., Mackin, R. B., Katsoyannis, P. G., Tycko, R., and Weiss, M. A. (2010) An Achilles' Heel in an amyloidogenic protein and its repair: insulin fibrillation and therapeutic design. *J. Biol. Chem.* **285**, 10806-10821
2. Jacoby, E., Hua, Q. X., Stern, A. S., Frank, B. H., and Weiss, M. A. (1996) Structure and dynamics of a protein assembly. ^{1}H-NMR studies of the 36 kDa R_6 insulin hexamer. *J. Mol. Biol.* **258**, 136-157
3. Glidden, M. D., Aldabbagh, K., Phillips, N. B., Carr, K., Chen, Y.-S., Whittaker, J., Phillips, M., Wickramasinghe, N. P., Rege, N., Swain, M., Peng, Y., Yang, Y., Lawrence, M. C., Yee, V. C., Ismail-Beigi, F., and Weiss, M. A. (2017) An ultra-stable single-chain insulin analog resists thermal inactivation and exhibits biological signaling duration equivalent to the native protein. *J. Biol. Chem.*, Submitted
4. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283-291
5. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) 1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR* **5**, 67-81
6. Schaffer, M. L., Deshayes, K., Nakamura, G., Sidhu, S., and Skelton, N. J. (2003) Complex with a phage display-derived peptide provides insight into the function of insulin-like growth factor I. *Biochemistry* **42**, 9324-9334
7. Torres, A. M., Forbes, B. E., Aplin, S. E., Wallace, J. C., Francise, G. L., and Norton, R. S. (1995) Solution structure of human insulin-like growthfactor II: relationship to receptor and binding protein interactions. *J. Mol. Biol.* **248**, 385-401
8. Dror, V., Nguyen, V., Walia, P., Kalynyak, T. B., Hill, J. A., and Johnson, J. D. (2007) Notch signalling suppresses apoptosis in adult human and mouse pancreatic islet cells. *Diabetologia* **50**, 2504-2515