

# A consensus view of protein dynamics

Manuel Rueda<sup>\*†</sup>, Carles Ferrer-Costa<sup>\*†</sup>, Tim Meyer<sup>\*†‡</sup>, Alberto Pérez<sup>\*†</sup>, Jordi Camps<sup>†§</sup>, Adam Hospital<sup>\*†§</sup>, Josep Lluís Gelpi<sup>\*†‡</sup>, and Modesto Orozco<sup>\*†§¶</sup>

<sup>\*</sup>Molecular Modelling and Bioinformatics Unit and <sup>§</sup>Structural Biology Node, Institut de Recerca Biomèdica, Parc Científic de Barcelona, Josep Samitier 1-5, 08028 Barcelona, Spain; <sup>†</sup>Computational Biology Program, Barcelona Supercomputing Center, Jordi Girona 31, Edifici Nexus II, 08028 Barcelona, Spain; and <sup>‡</sup>Departament de Bioquímica i Biologia Molecular, Facultat de Biologia, Universitat de Barcelona, Avda Diagonal 645, 08028 Barcelona, Spain

Edited by Harold A. Scheraga, Cornell University, Ithaca, NY, and approved November 10, 2006 (received for review July 6, 2006)

The dynamics of proteins in aqueous solution has been investigated through a massive approach based on “state of the art” molecular dynamics simulations performed for all protein metafolds using the four most popular force fields (OPLS, CHARMM, AMBER, and GROMOS). A detailed analysis of the massive database of trajectories (>1.5 terabytes of data obtained using ≈50 years of CPU) allowed us to obtain a robust-consensus picture of protein dynamics in aqueous solution.

force field | molecular dynamics | molecular modeling | protein structure

Flexibility is a key determinant of the biological functionality of proteins. A significant percentage of proteins are unfolded in the absence of ligand, and many others change their conformation as a result of the presence of other molecules or changes in the environment (1–4). Experimental representation of flexibility, even *a priori* possible, is in general difficult, which makes atomistic simulation the only viable alternative to study this important phenomenon for many proteins. Among the different theoretical methods available for description of protein flexibility, molecular dynamics is probably the most powerful (5–13). Since it was first applied to proteins in the late 1970s, molecular dynamics (MD) has been largely used to study the dynamics of proteins (14, 15). Unfortunately, due to the cost of simulations and the diversity of force fields, a consensus view of protein dynamics has not been yet obtained using this technique. In this paper, we present a systematic study of the most populated protein metafolds [see [supporting information \(SI\) Data Set](#)] using state of the art atomistic MD simulation conditions and the four most widely used force fields [OPLS (O), CHARMM (C), AMBER (A), and GROMOS (G)] (16–25). The result of this massive supercomputer effort (>1.5 terabytes of data and a computational equivalent to 50 CPU years) is a consensus picture of protein dynamics under conditions close to the physiological ones.

## Results and Discussion

**Supporting Information.** For further details, see [SI Data Set](#) and [SI Figs. 5–17](#)

**Force Field Convergence.** Previous to any dynamic study, we need to determine whether force fields are providing a similar picture of protein structure and whether such a picture is similar to that derived experimentally. Analysis of collected samplings indicates an average divergence ( $\alpha$  in Eq. 1) of ≈2 Å between force fields (slightly larger deviations are found in G simulations), which, considering the thermal noise of the simulations ( $\Omega$  in Eq. 2), suggests a similarity of ≈70% between the four samplings and an average “effective distance ( $\Omega^{-1}$ ) between them of only 1.4 Å (slightly larger values are always obtained for G-simulations; see Table 2). This finding indicates that all simulations are, in fact, sampling a similar region of the conformational space. This suggestion is supported by the analysis of the radii of gyration and solvent accessible surface (differences of ≈1% in radii of gyration and 2% in solvent-accessible surface between the four force fields; see Fig. 1).

Not only are the structures sampled in the four simulations similar, but they are also close to the experimental conformations

**Table 1. Structures representative of protein metafolds**

PDB ID code	Exp structure	No. of disulfide bridges	Atoms in simulation
1AGI	X-ray	3	23,403/22,734
1BFG	X-ray	0	25,212/24,509
1BJ7	X-ray	2	19,371/18,518
1BSN	NMR	0	33,753/32,948
1CHN	X-ray	0	18,022/17,282
1CQY	X-ray	0	28,856/28,310
1CSP	X-ray	0	13,293/12,949
1CZT	X-ray	1	31,072/30,167
1EMR	X-ray	0	32,808/31,885
1FAS	X-ray	4	15,709/15,398
1FVQ	NMR	0	16,322/15,921
1GND	X-ray	0	68,647/66,213
1I6F	NMR	4	15,837/15,552
1IL6	NMR	4	27,464/26,486
1J5D	NMR	0	23,184/22,682
1JLI	NMR	1	36,307/35,615
1K40	X-ray	0	37,686/36,883
1KTE	X-ray	1	19,490/18,869
1KXA	X-ray	0	24,969/24,113
1LIT	X-ray	3	26,289/25,651
1LKI	X-ray	3	33,388/32,403
1NSO	NMR	0	31,774/31,134
1OOI	X-ray	3	21,754/21,057
1OPC	X-ray	0	22,833/22,245
1PDO	X-ray	0	30,947/30,199
1PHT	X-ray	0	19,995/19,554
1SDF	NMR	2	32,579/32,168
1SP2	NMR	0	17,093/16,930
1SUR	X-ray	0	35,945/34,691
2HVM	X-ray	3	32,956/31,550

For each protein, we quote the number of disulfide bridges, the origin of experimental structure, and the number of atoms in the simulation box (A, C, O/G). PDB, Protein Data Bank.

(see Fig. 1). Thus, differences in radii of gyration between MD samplings and experimental structures are <1%, and average differences in solvent accessible surface are ≈5%. The average backbone rmsds between simulated and experimental structures are ≈2.0 Å (A, 1.9; C, 2.0; O, 1.9; G, 2.5 Å), close to the thermal noise of MD simulations. Quite surprisingly, the presence or absence of disulfide bridges does not modify the deviation of our samplings from experimental structures, which is, however, dependent on the origin of the experimental structure. Thus, proteins

Author contributions: M.R., C.F.-C., and T.M. contributed equally to this work; M.R., C.F.-C., A.H., J.L.G., and M.O. designed research; M.R., C.F.-C., T.M., A.P., and J.C. performed research; M.R., C.F.-C., T.M., and A.P. analyzed data; and M.O. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS direct submission.

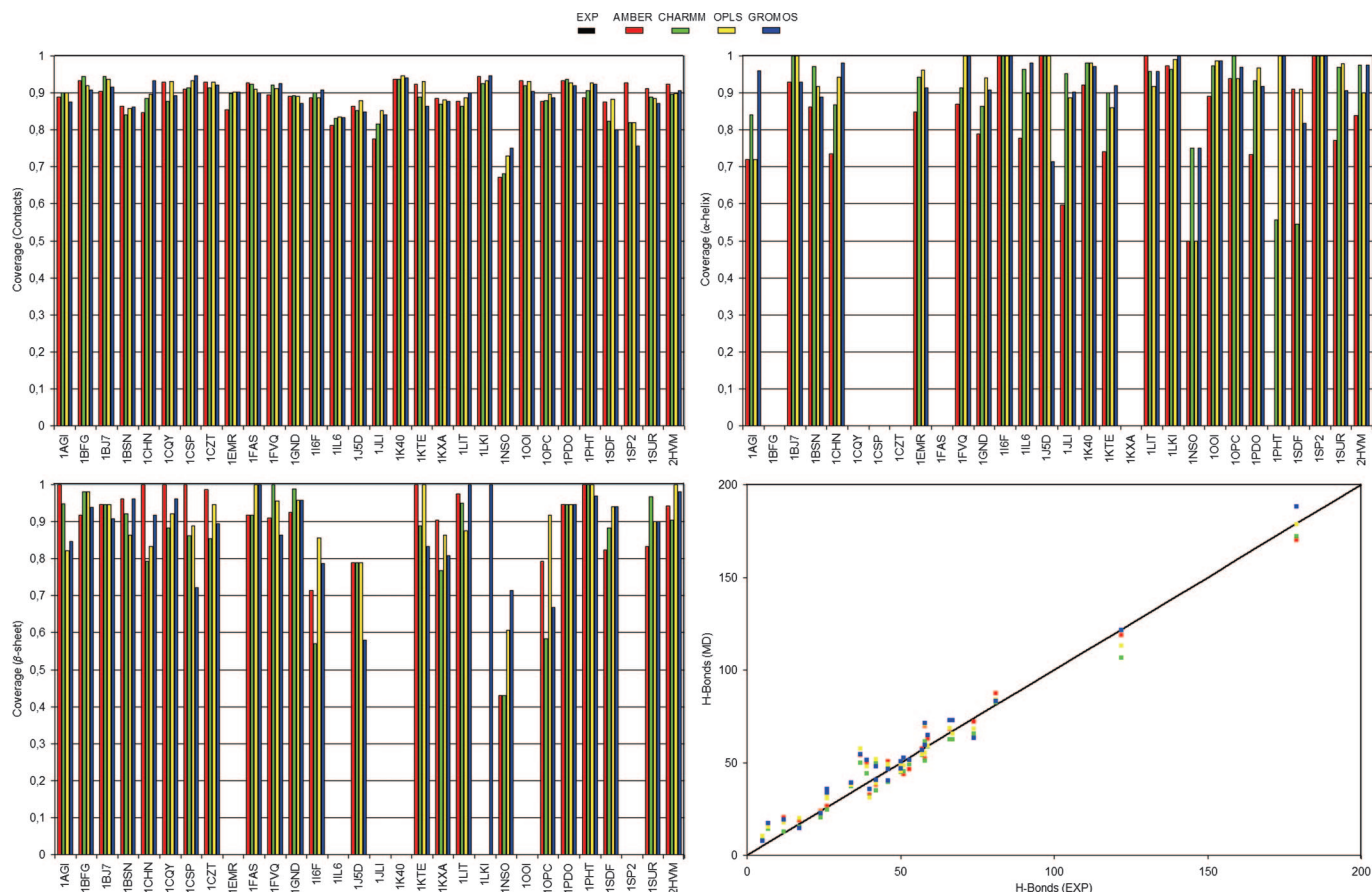
Abbreviations: MD, molecular dynamics; O, OPLS; C, CHARMM; A, AMBER; G, GROMOS.

<sup>¶</sup>To whom correspondence should be addressed. E-mail: [modesto@mmb.pcb.ub.es](mailto:modesto@mmb.pcb.ub.es).

This article contains supporting information online at [www.pnas.org/cgi/content/full/0605534104/DC1](http://www.pnas.org/cgi/content/full/0605534104/DC1).

© 2007 by The National Academy of Sciences of the USA





**Fig. 2.** Coverage of contacts (*Upper Left*), conservation of secondary structure for  $\alpha$ -helix (*Upper Right*),  $\beta$ -sheet (*Lower Left*), and relation between experimental and MD hydrogen bonds (intra; *Lower Right*) for A, C, O, and G force fields.

simulations are those with the largest mobility in the x-ray structures. It seems that MD-simulation is just allowing the movement of flexible residues that were “frozen” by the crystal lattice. This hypothesis can be tested by computation of crystal effective temperatures (see Eq. 3). Considering harmonic residue oscillations as those with B-factors below  $60 \text{ \AA}^2$  (see Fig. 3) we obtain an “effective temperature” of 290 K for the crystal. This demonstrates that the effect of crystal lattice is not, as often assumed, a general “effective cooling” of the general protein. The reduction in kinetic energy (i.e., the reduction of “effective temperature”) induced by the lattice is localized in a few very flexible residues, whose mobility is severely reduced with respect to the situation found in diluted aqueous solution.

Lindemann’s disorder index (see Eq. 4) provides a global picture of the flexibility of proteins compared with that of macroscopic solid or liquids (26). Quite independent of the force field and on the presence or absence of disulfide bridges MD simulations suggest that proteins as a whole behave like dense liquids ( $\Delta_L = 0.28 \pm 0.06$ ). However, this is the combination of the “solid-like” properties of the buried interior  $\Delta_L = 0.18 \pm 0.04$  ( $0.16 \pm 0.04$  if only buried backbones are considered) and of pure liquid properties of the exposed side-chains  $\Delta_L = 0.38 \pm 0.07$ . In summary, as suggested by experimental studies in particular systems (26, 27), proteins are “melted-solids” with strong differences between the near-solid interior and the full-liquid exterior.

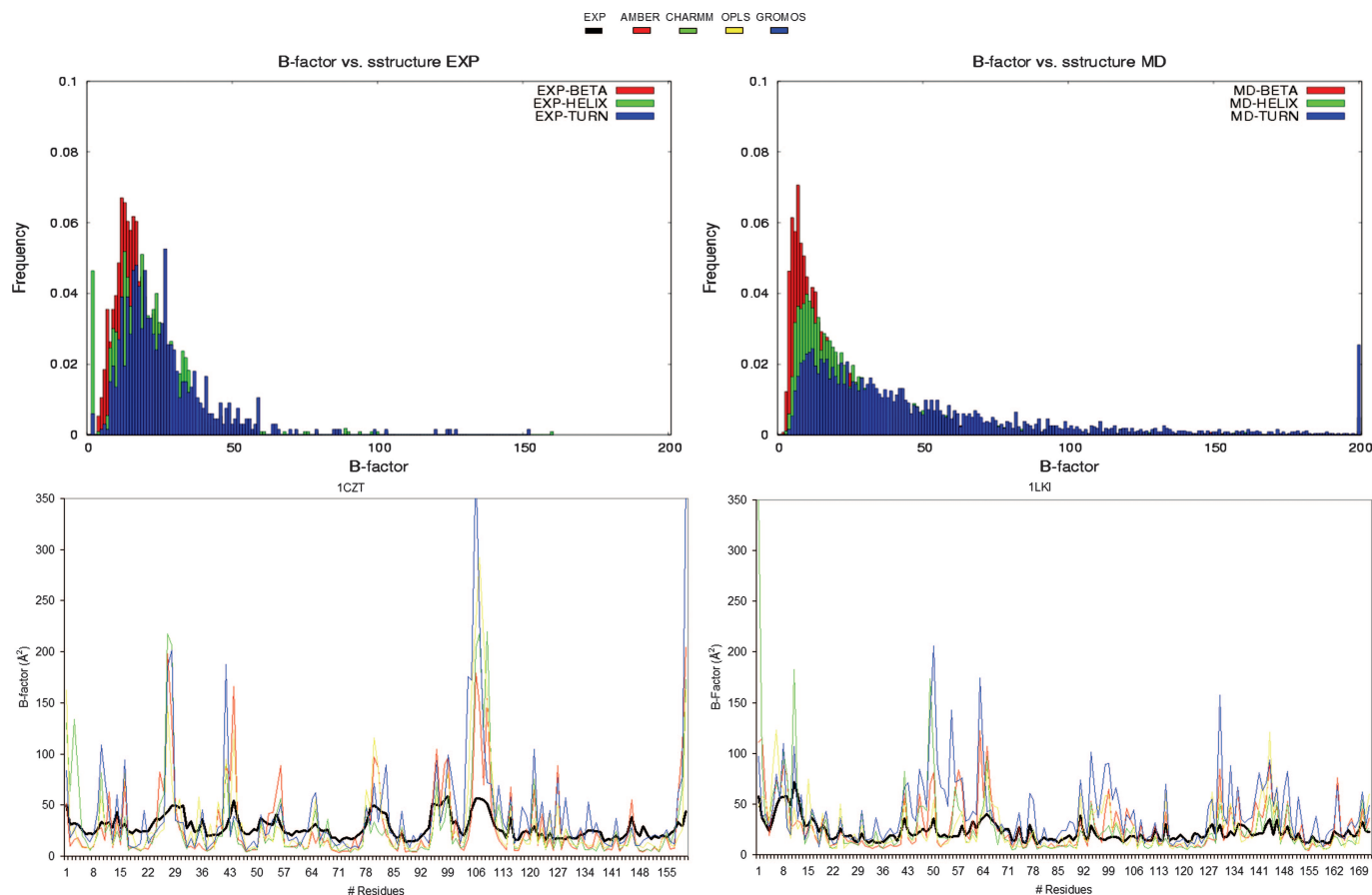
Essential dynamics provides an additional test of the global flexibility of proteins. Interestingly, all of the force fields provide a similar picture of the type of movements that are more important in describing protein similarity ( $\Gamma \approx 85\%$  see Table 2 and SI Figs. 6–15). Space dimensionality (i.e., the number of relevant deforma-

tion nodes, ref. 28) depends linearly with the number of residues of the protein ( $r^2 > 0.94$ ; see Fig. 4), with no clear changes induced by the presence of disulfide bridges. Given A, C, and O simulations, we can derive a consensus regression equation ( $\text{dim} = 32.5 + 0.58 \times \text{Nres}$ ;  $r^2 = 0.93$ ), which indicates that each residue adds only 0.6 dimensions to the conformational space, indicating that the global structure of the protein strongly limits, in a noncooperative manner, the number of accessible conformations of the constituting peptides.

The first essential deformation movements have associated very low frequencies ( $< 70 \text{ cm}^{-1}$  for the first 25 essential modes in the  $C_\alpha$  space) indicating that significant deformations are expected at room temperature along these modes. Entropies derived from these frequencies (see *Materials and Methods*) depend in all cases linearly with the number of residues ( $r^2 > 0.98$ ; see Fig. 4), confirming the lack of cooperative effects in determining the accessible conformational space of mono-domain proteins. Perfect match is found between A, O, and C simulations, whereas G trajectories lead to more disordered structures. Interestingly, the presence of disulfide bridges does not introduce any sizeable effect in determining the entropy of the folded protein. Thus, we can conclude that, in contrast with the general belief, disulfide bridges do not especially restrict the conformational space of native proteins.

A single experimental structure presents a clearly defined pattern of stabilizing interactions (salt bridges, hydrogen bonds, and hydrophobic interactions). Obviously, when flexibility is allowed, such a pattern becomes more diffuse and interacting groups are more promiscuous. Thus, despite the total number of stabilizing interactions preserved in MD, there is, in all cases, a reduction in the number of “permanent” interactions (defined as those occurring



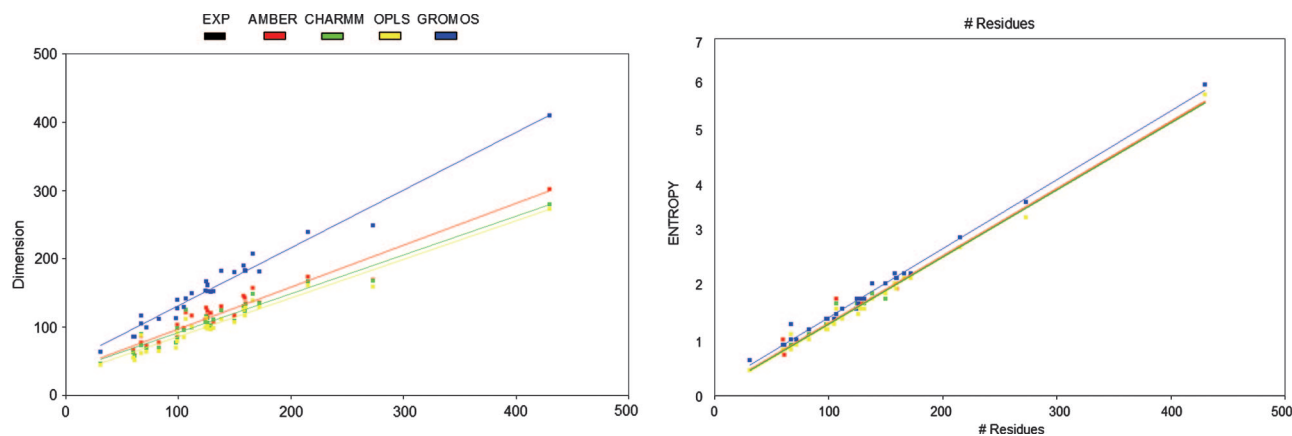


**Fig. 3.** B-factor distribution for x-ray structures and from MD simulations (*Upper*), and per residue B-factors for two proteins: 1CZT and 1LKI (*Lower*).

for >80% of the trajectory) with respect to those in the experimental structure. Thus, the number of permanent hydrogen bonds and hydrophobic interactions are  $\approx 68$  and 81% of those existing in experimental structures. The number of “permanent” salt bridges is  $\approx 77\%$  the experimental ones for A, C, and O simulations, whereas a bigger loss of salt bridges is found in G simulations (see Table 3). Once “permanent” interactions are located, we can determine their contribution to the global stiffness of the protein by deriving the associated force-constant (see Eq. 6). Results in Table 2 show that, in agreement with chemical intuition and irrespective of the force field used, hydrogen bonds are individually stiffer (force

constants in the range 22–29 kcal/mol  $\text{\AA}^2$ ) than hydrophobic contacts (force constants in the range 3–6 kcal/mol  $\text{\AA}^2$ ). More surprising is the soft nature of salt bridges (29), especially of those involving lysine, which in fact behave similarly (in terms of stiffness) to hydrophobic contacts. MD strongly suggests that individually, hydrogen bonds are the main responsible of protein stiffness, whereas salt bridges (especially with Lys as donor) are quite labile and promiscuous. It is then clear that none of the theoretically more stabilizing interactions disulfide and salt bridges are controlling protein dynamics.

Protein dynamics is complex; even in the hydrophobic core, side



**Fig. 4.** Dimensionality versus the number of residues (*Left*) and Schlitter's entropy in kcal/mol $^{-1}$ ·K $^{-1}$  (*Right*) for A, C, O, and G force fields.

**Table 3. Force constants (in kcal/mol Å<sup>2</sup>) associated with different types of protein interactions**

Interaction		AMBER	CHARMM	OPLS	GROMOS
Salt bridge	Total	11	15	17	—
	R-D	14	21	28	—
	R-E	15	20	21	—
	K-D	7	10	10	—
	K-E	5	8	8	—
H bonds	Total	29	26	26	22
	Back-back	27	24	24	22
	Others	35	31	33	23
Hydrophobic	Total	5	5	6	3
	Arom	7	7	7	4
	Aliphatic	5	5	5	3
	Cross	6	5	6	3

GROMOS does not lead to a harmonic representation of salt bridges, and accordingly, was not considered for this particular analysis.

chains may undergo rotameric changes on a time scale longer than that studied in the present work, which is limited to study short time scale dynamics. Furthermore, classic force fields are rough approximations to the real potential energy function. However, our massive dynamic analysis provides strong evidence that most of the behavior detected during MD simulations is force field independent and that deviations from experimental structures, when they exist, are not spurious, but signal a physically meaningful process. Proteins behave like biphasic systems, with a solid core surrounded by external liquid layer, whose flexibility is mostly restricted by hydrogen bonds, not by disulfide or salt bridges. At the monodomain level, the complexity of the accessible conformational space increases slowly, but linearly with the size of the protein, with no obvious cooperativity effects. Finally, and quite surprisingly, crystal lattice does not reduce the kinetic energy of all atoms of the protein, but just of a small set of otherwise too-flexible atoms, those that, when the lattice is released, lead to the largest conformational changes.

## Materials and Methods

**Model Selection.** We built a list of proteins to simulate the 30 most populated folds according to SCOP (30, 31), CATH (32), Dali (33, 34), and Dagget's databases (35) (see [SI Data Set](#)). When several good three-dimensional models exist in a class, we favor structures without bound ligands.

**System Setup.** As in our MODEL database (<http://mmb.pcb.ub.es/MODEL>), we used a common, automatic setup procedure designed to guarantee reasonable ionization states, no electrostatic unbalances, and a good hydration before simulation starts. Thus, in all cases, experimental structures (for NMR, the first one deposited in the Protein Data Bank) were titrated to define the major ionic state a neutral pH, neutralized by ions, minimized, heated, and hydrated (with special care in introducing structural waters). These systems were then preequilibrated for 0.5 ns with parm99-AMBER force field, and then equilibrated (0.5 ns) with each force field (see below).

**Simulation Details.** Equilibrated structures were used as starting points for 10-ns production trajectories, performed at constant pressure (1 atm) and temperature (300 K) using standard coupling schemes (the same in all cases). Trajectories for three (Protein Data Bank ID codes 1CQY, 1OPC, and 1KTE) ultrarepresentative structures of  $\alpha$ ,  $\beta$ , and  $\alpha/\beta$  folds (28) were extended to 100 (A, C, and O) or 50 (G) ns to evaluate the reliability of 10 ns trajectories. Results in [SI Fig. 16](#) strongly suggest that 10 ns is long enough for simulations for many analysis purposes. Particle Mesh Ewald approaches were used to deal with long-range effects (36). Integration of motion equations was performed every 1 fs, the vibrations of

bonds involving hydrogen being removed by SHAKE/RATTLE algorithm (37, 38). Four force fields were used in production runs: AMBER (17) (A), CHARMM (18, 19) (C), OPLS/AA (20–23) (O), and GROMOS (24, 25) (G). TIP3P (39, 40) was used as water model for A, C, and O simulations, whereas G calculations were done using the SPC model (41) for coherence with the force field. Simulations were done using parallel versions of AMBER8 (42) (A-simulations), NAMD (43, 44) (C and O simulations), and single processor version of GROMACS (45, 46) (G simulations).

Control simulations were performed to determine whether force fields were able to recognize unfolding conditions and lead to the destabilization of the folded form. For this purpose, 30-ns simulations of 1CQY, 1OPC, and 1KTE were performed using 8 M urea and a temperature of 368 K. OPLS (47) parameters for urea were used for O and A simulations, whereas specific CHARMM parameters (48) were used in C simulations. Due to the lack of specific parameters, these control simulations were not performed using GROMOS force field. Results in [SI Fig. 16](#) confirm that force fields are able to recognize strong denaturing conditions starting the unfolding of the proteins (which might take micro- to millisecond to complete) in the 0- to 30-ns simulation window. The final set with control simulations (10 ns; A, O, C, and G trajectories) was performed for three proteins (Protein Data Bank ID codes 2GB1, 1LYS, and 1UBQ) for which residual dipolar couplings and S<sup>2</sup> distributions have been reported (see references in [the legend of SI Fig. 17](#)) from NMR experiments. Comparison of MD- and NMR-derived measures of flexibility is quite satisfactory (see [SI Fig. 17](#)) giving confidence in the quality of our theoretical estimates of flexibility (it is not possible to perform this comparison for representative proteins in [SI Table 4](#) due to the lack of experimental data).

**Analysis of Trajectories.** Trajectories were analyzed to obtain structural and dynamic properties. Structural descriptors include backbone rmsd, TM score (49), radii of gyration, solvent accessible surface for all heavy atoms (50), secondary structure (51), intramolecular hydrogen bonds, contact maps, Ramachandran plots (52, 53), solvent contact profiles, and others. The global similarity between the structures collected with the different force fields was obtained by computing the pair-cross rmsd (i.e., the rmsd between all of the snapshots collected in the two trajectories; see Eq. 1), and a related similarity index (see Eq. 2).

$$\alpha_{AB} = \frac{1}{M_A M_B} \sum_{k=1}^{M_A} \sum_{l=1}^{M_B} \left( \frac{1}{N} \sum_{t=1}^{3N} (x_{At} - x_{Bt})^2 \right)^{1/2}, \quad [1]$$

where  $N$  is the number of atoms and  $M$  is the number of frames

$$\Omega_{AB} = \frac{\alpha_{AA} + \alpha_{BB}}{2\alpha_{AB}}. \quad [2]$$

Local dynamic properties are represented by the B factors, which are compared with x-ray values to obtain a measure of “crystal effective” temperatures (see Eq. 3). Global flexibility was represented by Lindemann's indexes (26) (see Eq. 4) and by essential dynamics protocols considering either Cartesian or mass-weighted covariance matrices (54), which provided us with measures as protein entropy (55–57) or the dimensionality of the conformational space (28). The similarity between the nature of the essential deformation patterns of two trajectories (of the same protein) was determined by  $\Gamma$  index (see Eq. 5; ref. 58), considering a small set of 25 eigenvectors (which represent  $\approx 80$ –85% of protein variance).

$$T_{crys} = T_{MD} \left\langle \frac{B_{factors}^{crys}}{B_{factors}^{MD}} \right\rangle \quad [3]$$

$$\Delta_L = \frac{\left( \sum_i \langle \Delta r_i^2 \rangle / N \right)^{1/2}}{a'}, \quad [4]$$

where  $a'$  is the most probable nonbonded near-neighbor distance,  $N$  is the number of atoms, and  $\Delta r_i$  stands for the fluctuation of atom  $i$  from its average position.

$$\Gamma_{AB} = 2 \frac{\sum_{j=1}^n \sum_{i=1}^n (v_i^A \cdot v_j^B)^2}{\left( \sum_{j=1}^n \sum_{i=1}^n (v_i^A \cdot v_j^A)^2 + \sum_{j=1}^n \sum_{i=1}^n (v_i^B \cdot v_j^B)^2 \right)}, \quad [5]$$

where  $n$  is the number of size of the important space ( $n = 25$  here) and  $v$  stands for one unitary eigenvector. Note that  $\Gamma = 1$  indicates that both trajectories are sampling the same type of essential movements, whereas  $\Gamma = 0$  means that they are orthogonal.

The flexibility related to key interactions (hydrogen bonds, hydrophobic contacts, and salt bridges) was computed by assuming

the harmonic oscillator model (see Eq. 6). Because the use of harmonic model implies a Gaussian distribution of distances, we computed force constant ( $K$  in Eq. 6) for “stable” interactions considered as those found in  $>80\%$  of the trajectory. A salt bridge was defined when the distance  $N_\epsilon(K)/C_\zeta(R) - C_\gamma(D)/C_\delta(E)$  was  $<6.5$  Å, hydrogen bonds were annotated by using standard PTRAJ criteria (42), and hydrophobic contacts were counted when two hydrophobic groups ( $A, V, P, F, M, I, L, W$ ), which are not neighbors in the sequence, have their  $C_\beta$  atoms separated by  $<10$  Å.

$$K = k_b T / \langle \Delta X^2 \rangle, \quad [6]$$

where  $k_b$  is the Boltzmann constant,  $T$  is the absolute temperature, and  $\Delta X$  is the oscillation in the interaction distance from average values.

We are indebted to Prof. F. J. Luque and Dr. P. Bernadó for helpful discussions and A. Caballero-Herrera for CHARMM urea parameters. T.M. and A.P. are fellows of the Spanish and Catalan Ministries of Science. This work was supported by the Spanish Ministry of Education and Science (BIO2006-01602 and GEN2003-20642-CO9-07). All calculations were carried out on *MareNostrum*, a supercomputer with  $\approx 4,524$  64-bit Myrinet-connected processors.

- Waldron TT, Murphy KP (2003) *Biochemistry* 42:5058–5064.
- Lee-Wei Y, Bahar I (2005) *Structure (London)* 13:893–904.
- Sacquin-Mora S, Lavery R (2006) *Biophys J* 90:2706–2717.
- Remy I, Wilson IA, Michnick SW (1999) *Science* 283:990–993.
- McCammon JA, Gelin BR, Karplus M (1977) *Nature* 267:585–590.
- Karplus M, McCammon JA (1986) *Sci Am* 254:42–51.
- Allen MP, Tildesley DJ (1989) *Computer Simulation of Liquids* (Clarendon, Oxford).
- Brooks CL, III, Karplus M, Pettitt BM (1987) *Proteins: A theoretical Perspective of Dynamics, Structure and Thermodynamics* (Cambridge Univ Press, Cambridge, UK).
- Warshel A (1976) *Nature* 260:679–683.
- Ponder JW, Case DA (2003) *Adv Protein Chem* 66:27–85.
- van Gunsteren WF, Berendsen HJ, Hermans J, Hol WG, Postma JP (1983) *Proc Natl Acad Sci USA* 80:4315–4319.
- Tirado-Rives J, Jorgensen WL (1991) *Biochemistry* 30:3864–3871.
- Tirado-Rives J, Jorgensen WL (1990) *J Am Chem Soc* 112:2773–2781.
- Karplus M, McCammon JA (2002) *Nat Struct Biol* 9:646–652.
- Karplus M, Kuriyan J (2005) *Proc Natl Acad Sci USA* 102:6679–6685.
- Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona G, Profeta S, Jr, Weiner P (1984) *J Am Chem Soc* 106:765–784.
- Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Jr, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) *J Am Chem Soc* 117:5179–5197.
- MacKerell AD, Jr, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, et al. (1998) *J Phys Chem B* 102:3586–3616.
- MacKerell AD, Jr, Wiorcikiewicz-Kuczera J, Karplus M (1995) *J Am Chem Soc* 117:11946–11975.
- Damm W, Frontera A, Tirado-Rives J, Jorgensen WL (1997) *J Comput Chem* 18:1955–1970.
- Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) *J Am Chem Soc* 118:11225–11236.
- Kaminski G, Duffy EM, Matsui T, Jorgensen WL (1994) *J Phys Chem* 98:13077–13082.
- Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL (2001) *J Phys Chem B* 105:6474–6487.
- Ott K-H, Meyer B (1996) *J Comput Chem* 17:1068–1084.
- Hermans J, Berendsen HJC, Van Gunsteren WF, Postma JPM (1984) *Biopolymers* 23:1513–1518.
- Zhou Y, Vitkup D, Karplus M (1999) *J Mol Biol* 285:1371–1375.
- Lindorff-Larsen K, Best RB, Depristo MA, Dobson CM, Vendruscolo M (2005) *Nature* 433:128–132.
- Perez A, Blas JR, Rueda M, López-Bes JM, de La Cruz X, Luque FJ, Orozco M (2005) *J Chem Theory Comput* 1:790–800.
- Strop P, Mayo SL (2000) *Biochemistry* 39:1251–1255.
- Andreeva A, Howorth D, Brenner SE, Hubbard TJ, Chothia C, Murzin AG (2004) *Nucleic Acids Res* 32:D226–D229.
- Murzin AG, Brenner SE, Hubbard T, Chothia C (1995) *J Mol Biol* 247:536–540.
- Pearl F, Todd A, Sillitoe I, Dibley M, Redfern O, Lewis T, Bennett C, Marsden R, Grant A, Lee D, et al. (2005) *Nucleic Acids Res* 33:D247–D251.
- Dietmann S, Park J, Notredame C, Heger A, Lappe M, Holm L (2001) *Nucleic Acids Res* 29:55–57.
- Holm L, Sander C (1995) *Trends Biochem Sci* 20:478–480.
- Day R, Beck DA, Armen RS, Daggett V (2003) *Protein Sci* 12:2150–2160.
- Darden TL, York D, Pedersen L (1993) *J Chem Phys* 98:10089–10092.
- Ryckaert JP, Ciccotti G, Berendsen HJC (1977) *J Comput Phys* 23:327–341.
- Andersen HC (1983) *J Comput Phys* 52:24–34.
- Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) *J Chem Phys* 79:926–935.
- Mahoney MW, Jorgensen WL (2000) *J Chem Phys* 112:8910–8922.
- Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J (1981) in *Intermolecular Forces*, ed Pullman B (Reidel, Dordrecht, The Netherlands), pp 331.
- Case DA, Pearlman DA, Caldwell JW, Cheatham TE, III, Ross WS, Simmerling CL, Darden TL, Marz KM, Stanton RV, Cheng AL, et al. (2004) AMBER 8 (University of California, San Francisco).
- Kale L, Skeel R, Bhandarkar M, Brunner R, Gursoy A, Krawetz N, Phillips J, Shinozaki A, Varadarajan K, Schulten K (1999) *J Comput Phys* 151:283–312.
- Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) *J Comput Chem* 26:1781–1802.
- Berendsen HJC, van der Spoel D, van Drunen R (1995) *Comput Phys Commun* 91:43–56.
- Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ (2005) *J Comput Chem* 26:1701–1718.
- Tirado-Rives J, Orozco M, Jorgensen WL (1997) *Biochemistry* 36:7313–7329.
- Caballero-Herrera A, Nilsson A (2006) *J Mol Struct* 758:139–148.
- Zhang Y, Skolnick J (2004) *Proteins* 57:702–710.
- Hubbard SJ, Thornton JM (1993) NACCESS (Dept of Biochemistry and Molecular Biology, University College, London).
- Kabsch W, Sander C (1983) *Biopolymers* 22:2577–2637.
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) *J Appl Crystallogr* 26:283–291.
- Morris AL, MacArthur MW, Hutchinson EG, Thornton JM (1992) *Proteins* 12:345–364.
- Amadei A, Linssen AB, Berendsen HJ (1993) *Proteins* 17:412–425.
- Harris SA, Gavathiotis E, Searle MS, Orozco M, Laughton CA (2001) *J Am Chem Soc* 123:12658–12663.
- Schlitter J (1993) *Chem Phys Lett* 215:617–621.
- Andricioaei I, Karplus M (2001) *J Chem Phys* 115:6289–6292.
- Hess B (2000) *Phys Rev E Stat Nonlin Soft Matter Phys* 62:8438–8448.