20.2.7. Assessment of the simulation procedure

Equation (20.2.3.1) is a reasonable representation of the energy function of proteins. This point is illustrated here with results from 0.8 ns molecular-dynamics simulations of hen egg-white lysozyme (PDB entry 1lzt, 1.97 Å resolution), bovine pancreas ribonuclease A (5rsa, 2.0 Å resolution), bovine α-lactalbumin (1hfz, 2.3 Å resolution) and trypsin (2ptn, 1.55 Å resolution). [The coordinates for bovine α-lactalbumin were kindly provided by K. R. Acharya prior to their publication (Pike *et al.*, 1996).]

The proteins were fully hydrated, and the simulations were calculated with the *CHARMM*22 force field, using truncated octahedron periodic boundary conditions. The four proteins were overlaid with bulk water from an equilibrated simulation, and a 10 ps trajectory was calculated for the rearrangement of the water molecules around the protein with fixed protein atoms. The number of water molecules (4000 to 6000) required to hydrate the protein varied with the protein size (123–223 residues) and shape, with at least four layers of water molecules between the peripheral protein atoms and the walls of the boxes. The simulations were performed at constant pressure and temperature ($T = 300$ K and $p = 1$ atm) using the extended-system algorithms (Hoover, 1985; Nose, 1984) implemented in *CHARMM*. The 300 K constant temperature was maintained by coupling to an external bath, with a coupling constant of 25 ps. The *SHAKE* algorithm (Ryckaert *et al.*, 1977) was used to constrain bond lengths between hydrogen atoms and heavy atoms, allowing for a time step of 2 fs in the integration of the equations of motion. A non-bonded cutoff of 12 Å was used for the Lennard–Jones potential calculation. The electrostatic forces and energies were computed using the PME method (Darden *et al.*, 1993; Essmann *et al.*, 1995). The PME charge grid spacing was 0.7 Å, and the charge grid was interpolated with the direct sum tolerance set to 4.0×10^{-6}. The non-bonded pair lists were updated every 10 steps. Structures for analysis were saved every 0.1 ps.

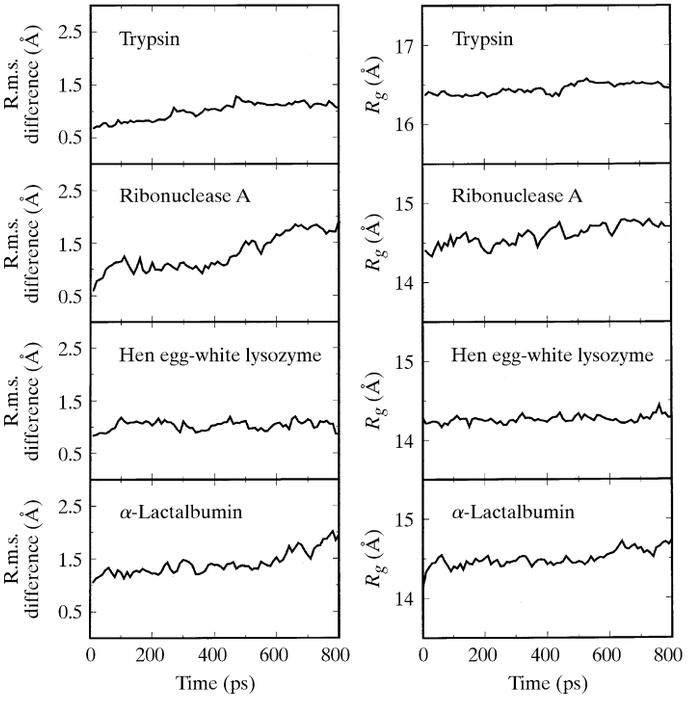

Fig. 20.2.7.1. Structural comparison and radii of gyration of various proteins as a function of time in the molecular-dynamics simulation. Left: r.m.s. coordinate differences averaged over main-chain atoms (N, Cα, C) between the energy-minimized crystallographic structure and the simulation snapshot. Right: radii of gyration (R_g).

Fig. 20.2.7.1 shows the deviation during the simulation period of the main-chain coordinates between the simulation structures and the crystallographic starting structure. The r.m.s. coordinate deviations in the case of lysozyme are particularly stable and small: approximately 1.0 Å over the course of the trajectory. Other r.m.s. values are more typical and range from 1.0 to 2.0 Å. The time dependence of other properties, such as the radius of gyration, can also be used to follow the stability and behaviour of a trajectory. These time series, also shown in Fig. 20.2.7.1, are constant. Jumps in such a time series can be used to detect conformational transitions (Post *et al.*, 1989).

Other experimental properties have been compared in the literature with those calculated from molecular-dynamics trajectories. Of particular interest is comparing time-dependent properties measured by NMR spectroscopy. An approach to calculating NMR relaxation rates was recognized early on when development of both the molecular-dynamics simulations of proteins and a model-independent theory for NMR relaxation was started (Levy, Karplus & McCammon, 1981; Levy, Karplus & Wolynes, 1981; Lipari & Szabo, 1982). Since then, the common practice of isotopic labelling of proteins for NMR structure determination has allowed the measurement of numerous NMR relaxation rates, particularly rates that characterize the motion of backbone atoms. Long simulations have been conducted to compare the calculated and experimental values (Abseher *et al.*, 1995; Chatfield *et al.*, 1998; Smith *et al.*, 1995). In a particularly long simulation, an 11 ns trajectory period was used to estimate relaxation rates associated with the motions of the vectors N—H, Cα—H and C—H methyl groups from alanine and leucine (Chatfield *et al.*, 1998). Trends in the general order of mobility of these vectors are reproduced, although a residue-by-residue comparison shows some differences.

20.2.8. Effect of crystallographic atomic resolution on structural stability during molecular dynamics

The variation in r.m.s. deviation between the initial crystallographic structure and the simulation coordinates for different protein trajectories (Fig. 20.2.7.1) raises the question of whether the atomic resolution of the starting X-ray structure influences the magnitude of this deviation. In order to investigate this issue, we calculated trajectories for bovine pancreatic trypsin inhibitor (BPTI), starting with crystallographic structures determined from data at three different atomic resolutions: 1bpi at 1.1 Å resolution (Parkin *et al.*, 1999), 6pti at 1.7 Å resolution (Wlodawer *et al.*, 1987) and 1bhc at 2.7 Å resolution (Hamiaux *et al.*, 1999). The errors in the atomic coordinates estimated from the Luzzati plots are 0.06 Å for the 1.1 Å resolution structure and 0.41 Å for the 2.7 Å resolution structure. The protocol described in the previous section was followed for simulations starting with each of the three crystallographic structures over a 500 ps simulation time. The net charge of +6 e on BPTI was neutralized by adding six chloride anions to the solvated protein system, thus accomplishing the ideal conditions for a PME calculation for the electrostatic interaction. The truncated octahedra contain approximately 3700 water molecules, and the total number of atoms in the simulations is over 12 000. The simulations were carried out on an eight-node IBM/SP2 and required 4.5 h of CPU time per 10 ps of dynamics run.

Root-mean-square differences (r.m.s.d.'s) in atomic coordinates were calculated between all pairs of coordinates from the X-ray structures, the energy-minimized X-ray structures and the 10 ps average MD structure obtained near 300 ps of the simulation period. In Table 20.2.8.1, the upper diagonal r.m.s.d. values are the main-chain-atom differences, while the lower diagonal ones are the side-chain-atom differences. The r.m.s.d.'s between the three X-ray structures range from 0.4–0.5 Å for the main-chain atoms and 1.4–

references