A Comparative Study of Carboxy Myoglobin in Saccharide−Water Systems by Molecular Dynamics Simulation

Grazia Cottone*

Dipartimento di Scienze Fisiche ed Astronomiche, Università Degli Studi di Palermo and CNISM, Via Archirafi 36, I-90123 Palermo, Italy

Received: November 21, 2006; In Final Form: January 29, 2007

Results from room-temperature molecular dynamics simulation on a system containing carboxy-myoglobin, water, and maltose molecules are reported. Protein atomic fluctuations, protein−solvent and solvent−solvent hydrogen bonding have been analyzed and compared to the ones in trehalose−water and sucrose−water systems (*Proteins* **2005**, *59*, 291−302). Results help in rationalizing, at a molecular level, the effects of homologues disaccharides on protein structure/dynamics experimentally observed. Furthermore, the effectiveness of disaccharides in bioprotection in terms of peculiar protein−matrix coupling is also discussed.

1. Introduction

Experimental studies performed on carboxy-myoglobin (MbCO) embedded in trehalose−water matrixes of low water content[1−9] showed that, in such systems, the internal protein dynamics is sizably reduced. Such reduction was also evidenced by molecular dynamics (MD) simulations of a MbCO molecule embedded in a plasticized amorphous trehalose matrix (89% trehalose/[trehalose + water] w/w, ∼2.3 water/sugar molar ratio).[10] Furthermore, results from simulation[11] evidenced that few trehalose molecules are bound to the protein mostly through single hydrogen bonds while water molecules bridge the protein surface and the sugar molecules. This is in agreement with the suggestion by Belton and Gil,[12] who proposed that trehalose, rather than directly binding to proteins, concentrates the residual water close to the macromolecules. Analogous results have been reported in a recent MD simulation of MbCO in a sucrose−water system[13] of identical composition (89% w/w) and in a simulation of lysozyme in a trehalose−water solution (18% w/w, ∼90 water/sugar).[14]

FTIR measurements, performed on samples of MbCO embedded in nonliquid water−trehalose matrixes at different hydration level,[15−19] gave information on the structural evolution of heme pocket and external matrix and evidenced the existence of a coupling between the internal degrees of freedom of the protein to those of the external matrix. Overall, the outcome of experiments and simulations suggest that the protein−solvent coupling arise from a hydrogen bond network involving protein groups, water and trehalose molecules, which anchors the protein surface to the surrounding medium,[15−20] and whose softness is modulated by the content of residual water (protein−water−sugar *structures*).

Recently, FTIR data on MbCO embedded in maltose−water,[19] compared with the previous ones on trehalose−water and sucrose−water,[17] pointed out that protein−matrix coupling depends in a subtle way on the detailed *solvent* composition. To rationalize at a molecular level the effects of homologues disaccharides on the protein structure/dynamics and the protein−solvent coupling, in this work, MD simulations are performed on MbCO embedded in a maltose−water box of the same composition (89% (sugar/[(sugar + water)] w/w, ∼2.3 water/maltose) as for trehalose and sucrose.[13] Protein atomic fluctuations, protein−solvent at the protein interface, and solvent−solvent hydrogen bonding have been analyzed; results are compared with the ones already reported for trehalose−water and sucrose−water systems of the same composition.[13]

On the basis of the phase diagram of sugar−water mixtures reported by Green and Angell,[21] the trehalose−water system is in a glassy state at room temperature, while sucrose−water and maltose−water are above the glass transition. Such differences in the physical state may have some implications on the structure and dynamics of the protein. It must be considered, however, that results relative to the physical state of the binary mixtures cannot be straightforwardly applied to protein-containing systems; indeed, FTIR data on MbCO−trehalose−water systems[15,18] point out that the presence of the protein causes a decrease in the matrix water content, suggesting that the protein competes for water with the sugar molecules thus affecting the hydrogen bond network within the sugar matrix in its proximity. Similar effect has been detected also in other sugars, showing a peculiar behavior in each sugar−water system (work in progress).

2. Computational Methods

MD simulations have been performed using the DL-PROTEIN program.[22] The Charmm22 force field has been adopted,[23] using for sugar a Charmm-type parameter set proposed for carbohydrates;[24] the TIP3P model was considered for the water molecules.[25]

Details on the simulations of the system containing MbCO, trehalose, or sucrose and water at 89% sugar/[(sugar + water)] w/w are described elsewhere;[13] these simulations were extended here to 3 ns. By following similar procedures, a system was built up containing one carboxy-myoglobin (PDB entry 1MB5[26]), maltose, and water, by using a pre-equilibrated maltose−water configuration as a building block, set up starting from the crystallographic coordinates of β-maltose monohydrate.[27] The protein was placed in a rectangular cell that was filled with maltose + water molecules. Molecules whose atoms were within a 1.8 Å distance from any protein atom were removed, thus leaving 260 maltose and 598 H_2O molecules, i.e., the same

* E-mail: cottone@fisica.unipa.it.

10.1021/jp0677288 CCC: $37.00
Published on Web 03/14/2007

sugar/water ratio as for the trehalose and sucrose systems (89% w/w, which correspond to ~2.3 water/sugar ratio).

Nine chlorine counterions were also included in the box to neutralize the total system charge by replacing nine water molecules near the box edges. The initial box length was ~60 × 58 × 60 Å^3; this ensures that almost all protein atoms are at least 10 Å away from the box edge.

A 200 ps MD simulation was first performed in the isobaric−isothermic ensemble[28] in order to equilibrate the system to 300 K and 1 atm, with coupling constants of 0.5 and 5 ps for the thermostat and the barostat, respectively. The system was then gradually heated from 300 to 600 K by increments of 50 K during 60 ps. Next, a 100 ps simulation was performed at 600 K in the NPT ensemble by coupling with a thermostat with time constant of 0.4 ps. The system was then gradually cooled to 300 K by decrements of 50 K during 60 ps; a 600 ps MD was then performed to equilibrate to 300 K and 1 atm.

The simulations for data collection were then performed in the microcanonical ensemble and consisted of a 3 ns trajectory at an average temperature of 303 ± 2. Periodic boundary conditions were used.[29] Van der Waals interactions have been cut off beyond a distance of 10 Å. Electrostatic interactions were calculated by the Ewald sum using the PME method.[30] The direct sum cutoff was 10 Å, the Ewald α parameter was set to 0.36, and eight order cubic spline interpolation with a grid of 55 × 50 × 55 points were used. All chemical bonds were kept fixed by using the SHAKE constraint algorithm.[31,32] The equations of motion were integrated with the extended velocity Verlet scheme[33,34] with a step size of 1 fs. Coordinate sets were saved every 200 fs for data analysis.

Information on protein internal motions is obtained through the analysis of atomic mean square fluctuations (MSFs), which are a measure of variability of the atomic position distribution functions. For the analysis of protein MSFs, each configuration was best fitted to the framework of the starting structure in order to correct for the effects due to possible rigid translations and rotations of the protein.[35] Letting the new configuration to be $\{r_i^*,\ i = 1,\ N\}$, the MSFs were calculated from the reconstructed configurations as the following isotropic average:

$$\frac{1}{3}[(\langle x_i^{*2}\rangle - \langle x_i^*\rangle^2) + (\langle y_i^{*2}\rangle - \langle y_i^*\rangle^2) + (\langle z_i^{*2}\rangle - \langle y_i^*\rangle^2)]\ i = 1, N$$

where the angular brackets indicate time averages. Following Maragliano et al.,[36] MSFs have been calculated by averaging over 15 time blocks of 200 ps length along the simulated trajectory of each protein−water−sugar system. Such a procedure gives unbiased measures of the variance of the distribution of atomic position, and it is particularly useful in the case of the multimodal distributions often encountered in proteins where, due to conformational changes, atoms can be delocalised among several sites.

The atomic fluctuations of the protein embedded in saccharide−water matrices were compared with the ones of a full hydrated carboxy-myoglobin at 300 and 400 K.[13] The high-temperature simulation in water enables understanding of which fluctuations along the main chain are responsible for protein thermal denaturation. This, in turn, allows understanding to which extent protein regions mainly involved in denaturation are protected by sugar coating and, more important, allows comparison of the efficiency of the different saccharides in bioprotection.

For the hydrogen bond analysis, we adopted the convention that a hydrogen bond is present when the donor−acceptor distance is shorter than 3.5 Å and the angle between the donor−acceptor axis and the donor−hydrogen bond is less than 30°.[37] Data on the hydrogen bond patterns were averaged over 15 time blocks of 200 ps length along each simulated trajectory of the protein−water−sugar systems.

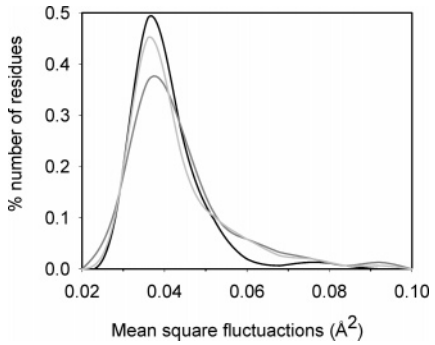

Figure 1. Distribution of atomic mean square fluctuations averaged over the protein main chain atoms (C, Cα, N): MbCO in 89% w/w trehalose−water (black line); MbCO in 89% w/w sucrose−water (dark-gray line); MbCO in 89% w/w maltose−water (gray line).

3. Results and Discussion

3.1. Mean Square Fluctuations. As mentioned when describing the computational methods, the protein mean square fluctuations (MSFs) were calculated by averaging over 15 time blocks of 200 ps along each simulated trajectory.[36]

MSFs values obtained averaging over the protein's main chain atoms are similar for MbCO in trehalose, sucrose, and maltose, being ~0.04 Å^2 at room temperature in each system, to be compared with 0.26 Å^2 obtained from a simulation in water at the same temperature. Such a behavior is in agreement with results from Mossbauer spectroscopy in 80% sucrose−water solution,[38] in dry trehalose-coated myoglobin,[5] and from neutron scattering on trehalose-coated MbCO samples.[6,8] Figure 1 shows the distribution of the MSFs values in the three systems. As evident in sucrose and maltose, values in the range ~0.05−0.1 Å^2 are populated to a larger extent than those in trehalose. Furthermore, a detailed comparison of atomic fluctuations along the chain (Figure 2, upper panel) point out that, in sucrose and maltose, the residues with larger flexibility are located at the E, F, and H helices as well as at the EF and GH loops. Vitkup et al.[39] performed a MD simulation of MbCO in water in which water molecules were kept at 180 K, while the embedded protein was maintained at 300 K; this mimicked a "high viscosity solvent" uniformly surrounding the protein, whose MSFs resulted of very low amplitude and uniform over the whole chain. The MSFs relative to the simulation in trehalose, shown in Figure 2, are most similar to the ones reported by Vitkup et al. in the case of "cold solvent". The similitude suggests, in line with FTIR results,[19] that MbCO experiences stronger and more uniform constraint from the external matrix in trehalose than in sucrose or maltose; this in turn may play a role in the bioprotection by sugars. The atomic fluctuations of the protein embedded in saccharide−water matrices are also compared with the ones of a fully hydrated carboxy-myoglobin at 300 and 400 K,[13] i.e., under conditions in which thermal denaturation takes place (see Figure 2, lower panel). In water, some protein regions appear more involved in denaturation (E, F, and G helix residues, EF and GH loops). As evident from Figure 2, motions in these parts of the protein are more hindered in trehalose than in sucrose or maltose.

3.2. Solvent Composition at the Protein Interface. Several studies have been performed on sugar−water biomolecule systems to understand the molecular mechanisms at the basis

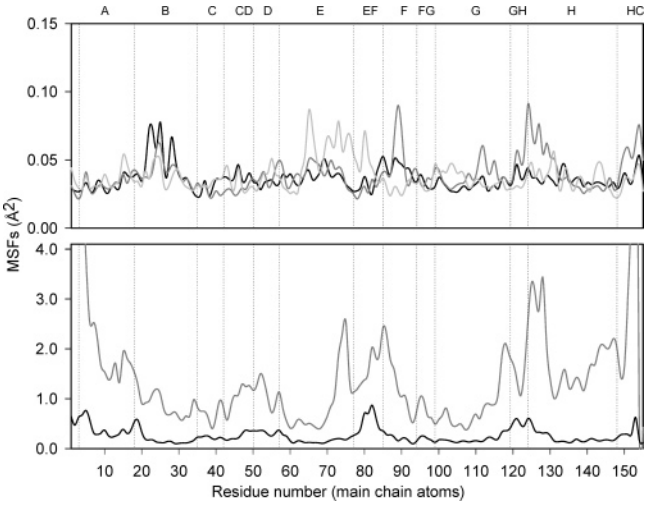

Figure 2. Atomic mean square fluctuations vs residue number, averaged over the protein main chain atoms (C, Cα, N). Upper panel: MbCO in 89% w/w trehalose−water (black line); MbCO in 89% w/w sucrose−water (dark-gray line); MbCO in 89% w/w maltose−water (gray line). Lower panel: MbCO in water at 300 K (black line); MbCO in water at 400 K (dark-gray line). The labels on the top of the plot indicate secondary structure elements in MbCO (see caption of Figure 2 in ref 10).

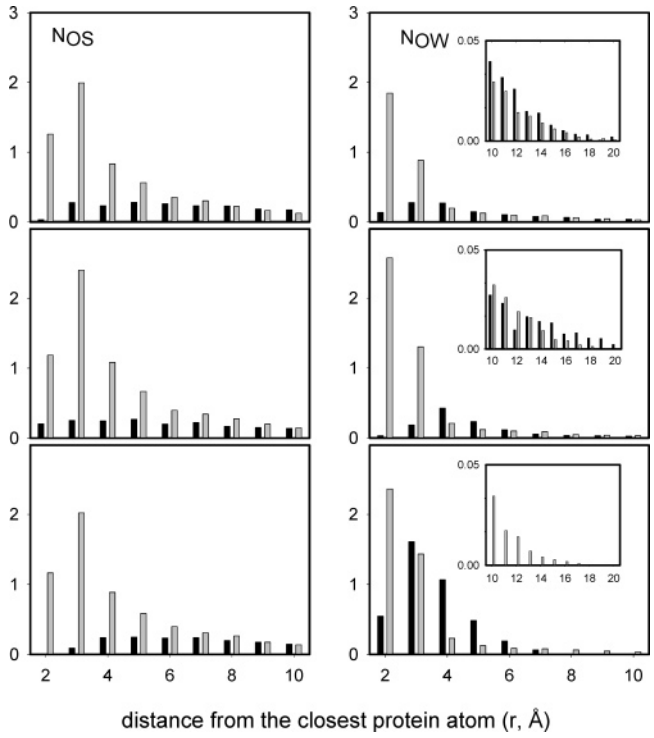

Figure 3. Number of water oxygen atoms N_{Ow} and number of oxygen atoms belonging to sugar hydroxyl (OH) groups N_{OS} found in class intervals 1 Å thick as a function of distance from the closest protein atom. Class i contains atoms whose distance from the closest protein atom lies between i and $i + 1$ Å. For each class, black bar: initial configuration, before the whole equilibration run; gray bar: average along the whole trajectory. Upper panels: trehalose−water system; central panels: maltose−water system; lower panels: sucrose−water system. Data are plotted only up to 10 Å from the closest protein atom; for distance larger than 10 Å (see insets in the right panels), the distributions trail off due to irregular edge effects in the rectangular box, lacking any physical meaning.

of trehalose peculiarity. The main hypotheses proposed are: (i) The *water replacement hypothesis*, according to which stabilization occurs via the formation of hydrogen bonds between the sugar and the biostructure;[40] (ii) The *water-entrapment hypothesis*, according to which, in the dry state, trehalose, rather than directly binding to proteins, traps the residual water at the biomolecule−sugar interface by glass formation;[12] (iii) *High viscosity hypothesis*, according to which viscosity effects cause motion inhibition[41] and hindering of processes which lead to loss of structure and denaturation. In this respect, Green and Angell[21] suggested the peculiarity of trehalose to be related to its rather high glass-transition temperature with respect to other glass-forming sugars.

Results above-discussed on MSFs, along with the experimental results, lead to suggest that bioprotection should be discussed in the framework of protein−solvent coupling. Then, to understand the effects of sugar matrices on the structure and dynamics of the protein, it is necessary to determine the distribution of the mixed solvent at the macromolecule interface. In this respect, atomic-resolution molecular dynamics is an alternative method to experiments.[42]

In this work, water and sugar molecules are classified, in the configurations generated by MD, according to the distance r of the water oxygen or of the sugar hydroxyl group oxygen from the closest protein atom. We recall that each sugar molecule contributes eight OH groups. Class intervals were chosen 1 Å thick. The distributions, averaged over the spherical surface area and over the whole simulation time, are reported in Figure 3 for the three systems. The analogous distributions in the initial configurations, before the equilibration runs, are also reported. Data are plotted only up to 10 Å from the closest protein atom; for distance larger than 10 Å (see the insets in the right panels), the distributions trail off due to edge effects in the rectangular box, lacking any physical meaning (see also Figure 5 in ref 14). These effects are due to the use of the minimum image convention in calculating atomic forces in the simulation. Indeed, the box half-length is ∼25 Å (see Methods section in this work and in ref 13), while the average gyration ratio of myoglobin is ∼15 Å; this leaves ∼10 Å distance from the protein surface and the box edges. Therefore, the layers for $r > 10$ Å are not fully contained within the minimum image.

A comparison between the initial configuration and the average over the trajectory evidence a relaxation of both water and sugar molecules, specific to each sugar−water systems, on the time scale of the simulation. This indicates that the results, discussed below, are not biased by initial condition effects.

Because of the high sugar concentration, it was not possible to have, in the nanosecond time scale, good enough statistics to obtain reliable values for the diffusion coefficients of both water and sugar molecules. Although the water mean square displacement, calculated along the trajectories as $\langle \Delta r(t)^2 \rangle$, does not reach a linear regime even at the longest time interval sampled, water translational diffusion is still observed, i.e., beyond the diameter of water molecule, over the time scale of our simulations. It is known that, in these systems, at high sugar concentration, water diffusion is decoupled from matrix viscosity;[43] indeed, open rigid three-dimensional hydrogen bond networks form in which water can still diffuse.[43−45] In this respect, the increase of the mean square displacement with time, shown in Figure 4, is the lowest in the maltose−water system. This result is in line with MD simulation results,[46] which suggested that maltose−water solutions are inhomogeneous systems in which maltose molecules cluster and water molecules remain trapped at these clusters. This argument has been used for giving a molecular basis to results obtained by FTIR measurements, which indicate that, in systems of water/sugar

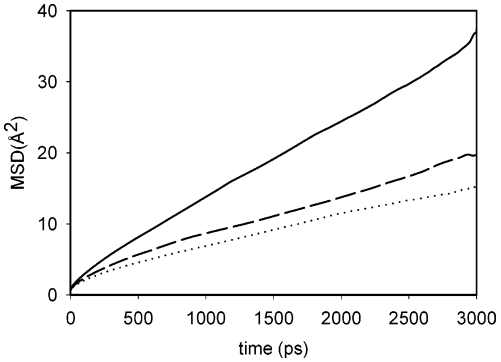

Figure 4. Mean square displacement vs time of water molecules in the trehalose−water system (solid line), sucrose−water system (dashed line), and maltose−water system (dotted line).

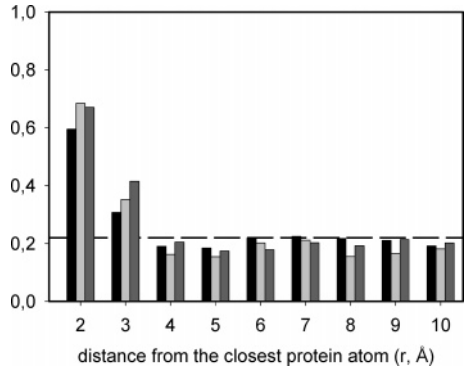

Figure 5. Histograms of the quantity $N_{Ow}/(N_{OS} + N_{Ow})$ for the sugar−water systems, averaged over the whole trajectory. Black bars: MbCO in 89% w/w trehalose−water; dark-gray bars: MbCO in 89% w/w sucrose−water; gray bars: MbCO in 89% w/w maltose−water. The dashed line indicates the average solvent composition (22%).

ratio ∼2−10, the matrix results are harder in maltose than in sucrose or trehalose.[19]

In Figure 5, we report the ratio $N_{OW}/(N_{OW} + N_{OS})$ for each sugar, where N_{OW} and N_{OS} indicate, respectively, the number of water oxygen atoms and sugar hydroxyl groups oxygen atoms found in each class interval, irrespective of their bound state. The second and third shells, which contain oxygen atoms whose distance from the protein is in the 1.0−3.0 Å interval (i.e., atoms which could exchange hydrogen bonds with the protein sites), are richer in water oxygen atoms than in sugar oxygen atoms. The value of the ratio $N_{OW}/(N_{OW} + N_{OS})$ in the second and third shells is sizably higher than the asymptotic value of 0.22 in each sugar system. Accordingly, we suggest that, even in systems at 89% w/w, water preferentially surrounds the protein. This is in agreement with the model proposed by Xie and Timasheff[47] for bioprotection in dilute sugar solutions or preferential hydration model, i.e., the solvent composition of the local protein domain is enriched in water relative to the average solvent composition in the outer shells ($r > 4.0$ Å). Similar results have been recently obtained from MD simulation of lysozyme in a trehalose−water solution at moderate sugar concentration (18% w/w [trehalose/(trehalose + water)], ∼90 water/trehalose molar ratio)[14] and further confirmed from recent experiments on MbCO/trehalose samples by Fayer and co-workers.[48]

3.3. Hydrogen Bond Analysis. The mean number of different water and sugar molecules hydrogen bonded to the protein is reported in Table 1 for the three systems. In Tables 2 and 3 is reported the distribution of the number of hydrogen bonds with the protein in which water or sugar OH groups are involved. These bonds are up to four for water, two as donor and two as acceptor, and up to three for each sugar hydroxyl group, one as donor and two as acceptor.

TABLE 1: Mean Number of Different Water and Sugar Molecules Hydrogen-Bonded to the Protein[a]

sugar/(sugar + water) w/w	water	sugar	OH = 1	OH = 2	OH = 3	OH = 4
89% trehalose	88 ± 2	55 ± 2	32 ± 2	17 ± 3 (58%)	4 ± 1 (31%)	0.9 ± 0.1 (7%)
89% sucrose	128 ± 3	49 ± 1	31 ± 2	13 ± 1 (63%)	5 ± 1 (26%)	0.1 ± 0.1 (10%)
89% maltose	140 ± 1	57 ± 1	33 ± 2	18 ± 1 (58%)	4 ± 1 (31%)	2.0 ± 1 (7%)

[a] In columns 2 and 3 is reported the average number of water and sugar molecules hydrogen bonded to the protein. Protein−solvent hydrogen bonds break and re-form on a picosecond time scale. Data are obtained by averaging over fifteen time-blocks of 200 ps length. In columns 4−7 is reported the average number of sugar molecules linked to the protein through n OH groups ($n = 1-4$); note that the number of OH groups per sugar molecule is 8. The percentages are relative to the total number of sugar molecules bound to the protein. The data in the table are irrespective of the number of hydrogen bonds performed by the bound water molecules or sugar OH groups.

TABLE 2: Mean Number of Different Water Molecules Forming *n* H-bonds with the Protein[a]

system	$n = 1$	$n = 2$	$n = 3$	$n = 4$	total
89% w/w trehalose	58 ± 2 (66%)	24 ± 1 (27%)	5 ± 1 (6%)	0.9 ± 0.4	88 ± 2
89% w/w sucrose	85 ± 3 (66%)	32 ± 1 (25%)	10 ± 1 (8%)	2 ± 1	128 ± 3
89% w/w maltose	95 ± 1 (68%)	36 ± 1 (26%)	8 ± 1 (6%)	0.7 ± 0.2	140 ± 1

[a] The percentages are relative to the total number of water molecules bound to the protein.

TABLE 3: Mean Number of Different Sugar OH Groups Forming *n* H-bonds with the Protein[a]

system	$n = 1$	$n = 2$	$n = 3$	$n = 4$	total
89% w/w trehalose	79 ± 1 (93%)	5 ± 1 (6%)	0.6 ± 0.1	0	85 ± 2
89% w/w sucrose	64 ± 2 (90%)	7 ± 1 (10%)	0.3 ± 0.1	0	71 ± 2
89% w/w maltose	80 ± 2 (89%)	9 ± 1 (10%)	0.5 ± 0.2	0	90 ± 2

[a] The percentages are relative to the total number of OH groups bound to the protein.

As shown in Table 1, trehalose, sucrose, and maltose molecules are bound to protein mainly through only one out of the eight OH groups available for binding (58%, 63%, 58% in trehalose, sucrose, and maltose, respectively). Tables 2 and 3 show that the protein exchanges either single or multiple hydrogen bonds with water, while the few sugar OH groups bound to the protein are in large majority linked through a single (donor) hydrogen bond. On the basis of the data shown in Table 1 and 3, one can conclude that the few sugar oxygen atoms found in the first hydration shell (see Figure 3) essentially belong to different sugar molecules. This notwithstanding the fact that, in each sugar molecule, the eight hydroxyl groups are topologically connected.

Results from simulations are in agreement with previous experimental data,[49] which suggested that the interaction of trehalose or sucrose with the protein is restricted to sites that are assumed to bind more mobile, singly hydrogen-bonded water. It was shown that the sugar−protein interaction does not affect the number of primary bound water molecules, which are likely to be tightly bound to the protein by multiple hydrogen

bonds.[49] Furthermore, along each simulated trajectory, the sugar molecules are found mainly bonded to the carboxylic oxygen atoms of glutamic and aspartic residues as well as to the peptide oxygen atoms, in agreement with IR data on trehalose−protein systems.[50]

Recent XAFS studies of cytochrome c in trehalose glasses at different content of residual water[51] pointed out a reduction of the mean square relative displacements of the central iron atom with respect the first coordination shell and a distortion of the porphirin group. The authors suggest that the matrix induces conformational changes in the protein, which in turn result in structural distortions of the heme group at the local atomic scale. Results from simulations show that a trehalose[11] and a maltose molecule are bound to the heme propionate oxygen atoms in the ∼90% and 76% of the total number of configurations saved along the trajectories, respectively; this suggests that the heme could be involved in a direct interaction with the matrix. At variance, a sucrose molecule was found bound to the heme propionate oxygen atoms only in the 15% of the total number of saved configurations.

As is well-known,[52] the dynamics of a protein is tightly coupled to the mobility of water molecules at the protein surface. It has been shown[52−54] that protein small-scale fluctuations are driven by hydrogen bond exchange between the protein surface and water molecules that undergo only librational/rotational motions. At variance, water translational motions, which allow complete exchange of protein-bound water molecules by translational displacement,[52−54] are necessary for large-scale fluctuations involving displacements of the protein surface. FTIR measurements[19] indicate that this behavior takes place in humid samples (∼20 water/sugar) at $T > \sim 200$ K, while only low tier, internal processes are present in very dry (∼0.2 water/sugar) and dry (∼2 water/sugar) samples even at room temperature.

The protein−sugar−water systems simulated are mostly comparable to the humid samples studied by FTIR experiments;[19] in such systems, the matrix thermal structural rearrangements are found almost overlapping in the three disaccharides, while differences are evident in the heme pocket thermal behavior. Such differences could be ascribed to the behavior of water at the interface as depending on the number of hydrogen bonding in which each molecule is involved. The different coupling in the different sugar matrixes can be rationalized based on data reported in Tables 2−3. The protein makes ∼220, ∼270, and ∼290 hydrogen bonds, with the water−sugar solution in trehalose, sucrose, and maltose, respectively; this might imply a more robust protein−water−sugar hydrogen bond network in maltose or sucrose with respect to trehalose. The above differences are essentially brought about by a larger number of water molecules close to the protein in sucrose or maltose with respect to trehalose. As Figure 3 shows, while the sugar atom distributions are nearly the same in the different systems, a larger number of water atoms is present in the second class interval, i.e., in the range 2−3 Å from the closest protein atom in sucrose and in maltose with respect to trehalose. However, the main contribution to this water excess in sucrose and maltose comes from the subset of water molecules bound to the protein through single hydrogen bonds, i.e., water molecules involved in relatively weaker interactions with the protein (see Table 2). A lower difference is found between the fraction of multiple bound water molecules than that between the number of bound sugar atoms (see Table 3).

Furthermore, there is a number of protein-bound water molecules, which is also bound to the sugar: 95, 84, 64 in maltose, sucrose, and trehalose systems, respectively (see Table 4). This give a fraction of bridging water molecules, over the total number of protein-bound water, which is higher in trehalose (73%) than in sucrose (66%) or maltose (68%). At variance, the distribution of the bridging water over the number of hydrogen bonds (columns 4−6 of Table 4) is nearly the same in the three systems (the differences being within the statistical error, estimated by block-averaging). Even for the shared water, the water excess in sucrose or maltose does not shift the distribution toward multiple hydrogen bonds, i.e., a larger number of water molecules could not imply a more rigid hydrogen bond network at the protein interface.

TABLE 4: Mean Number of Water Molecules, Shared between Sugar Molecules and the Protein, Forming *n* H-bonds with Protein and *m* with Sugar[a]

system	shared water		$n = 1$	$n = 2$	$n = 3$
89% w/w trehalose	64 ± 1 out of 88 ± 1 bound to protein (73%)	with protein	48 ± 2 (75%)	14 ± 1 (22%)	1 ± 1 (1%)
		with trehalose	37 ± 1 (58%)	22 ± 1 (34%)	5 ± 1 (8%)
89% w/w sucrose	84 ± 2 out of 128 ± 3 bound to protein (66%)	with protein	66 ± 2 (74%)	16 ± 1 (19%)	1 ± 1 (1%)
		with sucrose	51 ± 2 (61%)	28 ± 1 (33%)	4 ± 1 (5%)
89% w/w maltose	95 ± 1 out of 140 ± 1 bound to protein (68%)	with protein	72 ± 1 (76%)	20 ± 1 (21%)	3 ± 1 (3%)
		with maltose	57 ± 1 (60%)	32 ± 1 (34%)	6 ± 1 (6%)

[a] The percentages in columns 4−6 are relative to the total number of shared water molecules.

TABLE 5: Total Number of H-Bonds between Protein Secondary Structure Elements (Helices and Loops) and Water Molecules (Acting as Hydrogen Bond Donor)[a]

	secondary structure elements[b]											
system	A	B	C	CD	D	E	EF	F	FG	G	GH	H
89% w/w trehalose	72	37	68	26	23	53	22	25	20	41	30	62
89% w/w sucrose	91	72	32	27	48	59	38	25	14	57	46	92
89% w/w maltose	104	76	72	42	38	66	37	36	13	54	42	97

[a] Data refer to the 2.8−3.0 ns time slice of each trajectory. [b] See caption of Figure 2 in ref 10.

A *site specific* analysis has been also performed on the last 200 ps time slice by counting the hydrogen bonds formed by each protein residue and the water molecules. The analysis was restricted to water acting as hydrogen bond donor toward the protein atoms, which mostly contributes to the differences found in the three systems (the hydrogen bond pattern arising from sugar atoms is quite the same in each sugar−water system, data not shown). Furthermore, in this analysis, all of the hydrogen bonds are included, even when formed by the same water molecule; this give a total number of bonds that is larger than the one estimated from data reported in Table 2. Results reported in Table 5 for each protein secondary structure element show no clear-cut correlation between residue fluctuactions and the overall number of protein−residue hydrogen bonds. In particular, the number of bonds to the F helix is the same in trehalose and sucrose (in the latter system, the F helix has been found more flexible than in the former, see Figure 2), while the number of bonds to the helices G, H, and loop GH is quite the same in

maltose and sucrose (in the latter system, these protein region show the largest mobility among the three sugar systems). All of these results are in line with results obtained by FTIR spectroscopy on carboxy-myoglobin embedded in sugar/water matrices,[19] pointing out how the dynamics of both the external matrix and the protein is sizably reduced when lowering the content of residual water in the samples.

The above considerations suggest that only the fraction of water molecules shared between protein and sugar determine the strength of the constraints imposed by the matrix on the protein; then, this fraction could rationalize the larger local freedom for the protein internal motions in sucrose and maltose with respect to trehalose.

The fraction of water molecules bridging different sugar molecules, irrespective of their interaction with the protein, is 70%, 66%, and 63% for trehalose, maltose, and sucrose, respectively; this suggests that trehalose is more able to form *structures* in which water and sugar molecules cross-connect through the whole system,[15] i.e., even far from the protein surface.

The result is in line with simulative and experimental data on binary mixtures, indicating that the presence of trehalose sizably modifies the hydrogen bond network and the water dynamics by tightly binding water molecules.[55,56] In particular, a comparison with sucrose points out that, due to the formation of internal hydrogen bonds, sucrose is more rigid and less hydrated than trehalose because internal bonds leave a lower number of sites available for interactions with water;[57] maltose, notwithstanding the similarity in hydration number, restrains the surrounding water molecules less than trehalose.[58]

4. Conclusions

In this work, results from room-temperature molecular dynamics simulations on carboxy-myoglobin embedded in plasticized amorphous systems containing water and trehalose, sucrose, or maltose molecules are reported and compared. In agreement with Mossbauer[38,5] and neutron scattering data[6,8] on sugar-coated MbCO, the simulations indicated an overall reduction of protein atomic motions with respect to aqueous solutions in each sugar−water system. A comparison between the sugars points out that some regions of the protein chain exhibit a larger flexibility in sucrose and maltose than in trehalose. Such regions are plausibly involved in the thermally induced heme pocket structural variations detected by FTIR experiments on saccharide-coated myoglobin, recently reported.[17−19]

Further, on the nanosecond time scale, the sugars appear preferentially excluded from the protein domain. In all simulated systems, *structures* are formed in which the protein is confined within a network of hydrogen bonds connecting groups at the protein surface, water, and trehalose molecules. In such *structures*, water is in excess with respect to the average solvent composition, while relatively few sugar molecules are bound to the protein, mainly through single hydrogen bonds.

Result from simulations show that, in spite of a larger number of water molecules hydrogen-bonded to the protein in sucrose and maltose, the fraction of water molecules bound both to the protein and the sugar is lower in these systems with respect to the trehalose−water system. Assuming that the protein−matrix coupling is mediated by the subset of water molecules bridging the protein to the sugar, results from simulations suggest that the coupling is tighter in trehalose than in maltose and sucrose.

Furthermore, although the calculated water diffusion is overall lower in maltose and sucrose than in trehalose, protein mobility along the main chain is larger in the former systems. This is in agreement with FTIR results,[19] which indicate that the protein−solvent coupling is mostly regulated by the detailed composition of solvent at the protein surface, while barely depending on the water behavior far from the protein domain.

The protein−sugar−water systems simulated are mostly comparable to the humid samples studied by FTIR experiments.[19] This notwithstanding, results from simulations could shed light on the problem of bioprotection by saccharides in the limit of very low water content by suggesting that the ability of the sugar to protect biomolecules arises from the ability to lock the protein surface through a constrained, thin water layer, which hinders large-scale, solvent-coupled protein motions, leaving only some local motions that occur without displacements of the protein surface.[4,9] This interpretation has been confirmed recently also by vibrational echo experiments on MbCO embedded in trehalose glasses and silica gels.[48,59] The presence of tightly bound water at the protein−sugar interface could also give a microscopic basis for the increased water surface tension (i.e., the free energy cost to remove water from the surface) induced by cosolvent in ternary systems such as protein−water−sugar matrixes.[60]

Acknowledgment. I thank L.Cordone for reading of the manuscript, discussion, and criticisms, and G. Ciccotti for useful discussions. This work was supported by MIUR (grant PRIN 2005, Proprietà Dinamiche Strutturali e Funzionali di Proteine in Sistemi Non-Liquidi Contenenti Acqua Residua: Accoppiamento con la Matrice Esterna) and local fundings (ex 60%).

References and Notes

(1) Hagen, S. J.; Hofrichter, J.; Eaton, W. A. *Science* **1995**, *269*, 959−962.

(2) Gottfried, D. S.; Peterson, E. S.; Sheikh, A. G.; Wang, J.; Yang, M.; Friedman, J. M. *J. Phys. Chem.* **1996**, *100*, 12034−12042.

(3) Kleinert, T.; Doster, W.; Leyser, H.; Petry, W.; Schwarz, V.; Settles, M. *Biochemistry* **1998**, *37*, 717−733.

(4) Rector, D.; Jiang, J.; Berg, M. A.; Fayer, M. D. *J. Phys. Chem. B* **2001**, *105*, 1081−1092.

(5) Cordone, L.; Galajda, P.; Vitrano, E.; Gassmann, A.; Ostermann, A.; Parak, F. *Eur. Biophys. J.* **1998**, *27*, 173−176.

(6) Cordone, L.; Ferrand, M.; Vitrano, E.; Zaccai, G. *Biophys. J.* **1999**, *76*, 1043−1047.

(7) Librizzi, F.; Vitrano, E.; Cordone, L. *Biophys. J.* **1999**, *76*, 2727−2734.

(8) Librizzi, F.; Paciaroni, A.; Pfister, C.; Vitrano, E.; Zaccai, G.; Cordone, L. In *Proceedings of ILL Millenium Symposium and European User Meeting*; ILL: Grenoble: France, 2001; pp 58−59.

(9) Librizzi, F.; Viappiani, C.; Abbruzzetti, S.; Cordone, L. *J. Chem. Phys.* **2002**, *116*, 1193−1200.

(10) Cottone, G.; Cordone, L.; Ciccotti, G. *Biophys. J.* **2001**, *80*, 931−938.

(11) Cottone, G.; Ciccotti, G.; Cordone, L. *J. Chem. Phys.* **2002**, *117*, 9862−9866.

(12) Belton, P. S.; Gil, A. M. *Biopolymers* **1994**, *34*, 957−961.

(13) Cottone, G.; Giuffrida, S.; Ciccotti, G.; Cordone, L. *Proteins: Struct., Funct., Bioinf.* **2005**, *59*, 291−302.

(14) Lins, R. D.; Pereira, C. S.; Hunenberger, P. H. *Proteins: Struct., Funct., Bioinf.* **2004**, *55*, 177−186.

(15) Giuffrida, S.; Cottone, G.; Librizzi, F.; Cordone, L. *J. Phys. Chem. B* **2003**, *107*, 13211−13217.

(16) Giuffrida, S.; Cottone, G.; Librizzi, F.; Cordone, L. In *Progress in Condensed Matter Physics: Festschrift in Honour of Vincenzo Grasso*; Mondio, G., Sipigni, L., Eds.; Italian Physical Society: Bologna, Italy, 2004; pp 131−138.

(17) Giuffrida, S.; Cottone, G.; Cordone, L. *J. Phys. Chem. B* **2004**, *108*, 15415−15421.

(18) Cordone, L.; Cottone, G.; Giuffrida, S.; Palazzo, G.; Venturoli, G.; Viappiani, C. *Biochim. Biophys. Acta* **2005**, *1749*, 252−281.

(19) Giuffrida, S.; Cottone, G.; Cordone, L. *Biophys. J.* **2006**, *91*, 968−980.

(20) Abbruzzetti, S.; Giuffrida, S.; Sottini, S.; Viappiani, C.; Cordone, L. *Cell Biochim. Biophys.* **2005**, *42*, 431−437.

(21) Green, J. L.; Angell, C. A. *J. Phys. Chem.* **1989**, *93*, 2880−2882.

(22) Melchionna, S.; Cozzini, S. *DL-PROTEIN User Manual*; http://www.sissa.it/cm/DLPROTEIN. DL-PROTEIN has been developed from DL_POLY for the simulation of proteins by the Istituto Nazionale di Fisica della Materia, under the Network on "MD simulation of biosystems", group University of Rome La Sapienza. DL-POLY is a package of molecular simulation routines written by W. Smith and T. R. Forrester, copyright The Council for Central Laboratory of the Research Councils; Daresbury Laboratory at Daresbury: Nr. Warrington, U.K., 1996.

(23) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.

(24) Ha, A. S. N.; Giammona, A.; Field, M.; Brady, J. W. *Carbohydr. Res.* **1988**, *180*, 207−221.

(25) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *9*, 26−935.

(26) Cheng, X.; Schoenborn, B. P. *Acta Crystallogr., Sect. B: Struct. Sci.* **1990**, *46*, 195−208.

(27) Gress, S. M. E.; Jeffrey, G. A. *Acta Crystallogr., Sect. B: Struct. Sci.* **1977**, *33*, 2490−2495.

(28) Melchionna, S.; Ciccotti, G. *J. Chem. Phys.* **1997**, *106*, 195−199.

(29) Allen M. P.; Tildesley D. J. *Computer Simulation of Liquids*; Clarendon Press: Oxford, 1987.

(30) Essmann, U.; Perera, L.; Berkowitz, M.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(31) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.

(32) Ciccotti, G.; Ryckaert, J. P. *Comp. Phys. Rep.* **1986**, *4*, 345−392.

(33) Andersen, H. C. *J. Comput. Phys.* **1983**, *52*, 24−34.

(34) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. *Mol. Phys.* **1996**, *87*, 1117−1157.

(35) Kneller, G. R. *Mol. Simul.* **1991**, *7*, 113−119.

(36) Maragliano, L.; Cottone, G.; Cordone, L.; Ciccotti, G. *Biophys. J.* **2004**, *86*, 2765−2772.

(37) Luzar, A.; Chandler, D. *Phys. Rev. Lett.* **1996**, *76*, 928−930.

(38) Lichtenegger, H.; Doster, W.; Kleinert, T.; Birk, A.; Sepiol, B.; Vogl, G. *Biophys. J.* **1999**, *76*, 414−422.

(39) Vitkup, D.; Ringe, D.; Petsko, G. A.; Karplus, M. *Nat. Struct. Biol.* **2000**, *7*, 34−38.

(40) Carpenter, J. F.; Crowe, J. H. *Biochemistry* **1989**, *28*, 3916−3922.

(41) Sampedro, J. G.; Uribe, S. *Mol. Cell Biochem.* **2004**, *256*, 319−327.

(42) Karen, E.; Tang, S.; Bloomfield, V. A. *Biophys. J.* **2002**, *82*, 2876−2891.

(43) Rampp, M.; Buttersack, C.; Ludemann, H. D. *Carbohydr. Res.* **2000**, *328*, 561−572.

(44) Roberts, C. J.; Debenedetti, P. G. *J. Phys. Chem. B* **1999**, *103*, 7308−7318.

(45) Molinero, V.; Cagin, T.; Goddard, W. A., III. *Chem. Phys Lett.* **2003**, *377*, 469−474.

(46) Lerbret, A.; Bordat, P.; Affouard, F.; Descamps, M.; Migliardo, F. *J. Phys. Chem. B* **2005**, *109*, 11046−11057.

(47) Xie, G. F.; Timasheff, S. N. *Biophys. Chem.* **1997**, *64*, 25−43.

(48) Massari, A. M.; Finkestein, I. J.; McClain, B. L.; Goj, A.; Wen, X.; Bren, K. L.; Loring, R. F.; Fayer, M. D. *J. Am. Chem. Soc.* **2005**, *127*, 14279−14289.

(49) Lopez-Diez, E. C.; Bone, S. *Phys. Med. Biol.* **2000**, *45*, 3577−3588.

(50) Allison, S. D.; Chang, B.; Randolph, T. W.; Carpenter, J. F. *Arch. Biochem. Biophys*. **1999**, *365*, 289−298.

(51) Giachini, L.; Francia, F.; Cordone, L.; Boscherini, F.; Venturoli, G. *Biophys. J.* **2006**, *92*, 1350−1360.

(52) Doster, W.; Bacheitner, A.; Dunau, R.; Hiebl, M.; Luscher, E. *Biophys. J.* **1986**, *50*, 213−219.

(53) Tarek, M.; Tobias, D. J. *Biophys. J.* **2000**, *79*, 3244−3257.

(54) Tarek, M.; Tobias, D. J. *Phys. Rev. Lett.* **2002**, *88*, 8101−8104.

(55) Magazù, S.; Maisano, G.; Migliardo, P.; Mondelli, C. *Biophys. J.* **2004**, *86*, 3241−3249.

(56) Conrad, P. B.; de Pablo, J. J. *J. Phys. Chem. A* **1999**, *103*, 4049−4055.

(57) Ekdawi-Sever, N. C.; Conrad, P. B.; de Pablo, J. J. *J. Phys. Chem. A* **2001**, *105*, 734−742.

(58) Sakurai, M.; Murata, M.; Inoue, Y.; Hino, A.; Kobayashi, S. *Bull. Chem. Soc. Jpn.* **1997**, *70*, 847−858.

(59) Massari, A. M.; Finkelstein, I. J.; Fayer, M. D. *J. Am. Chem. Soc.* **2006**, *128*, 3990−3997.

(60) Timasheff, S. N. *Annu. Rev. Biophys. Struct.* **1993**, *22*, 67−97.