Modelling of the Complex between a 15-Residue Peptide from mSos2 and the N-Terminal SH3 Domain of Grb2 by Molecular-Dynamics Simulation

by Sofia Calero a), Santiago Lago b), Wilfred F. van Gunsteren c), and Xavier Daura d)

a) Department of Chemical Engineering, Faculty of Science, University of Amsterdam, NL-1018 WV Amsterdam
b) Department of Environmental Sciences, Faculty of Experimental Sciences, Universidad Pablo de Olavide, ES-41013 Sevilla
c) Laboratory of Physical Chemistry, Swiss Federal Institute of Technology Zurich, ETH Hönggerberg, CH-8093 Zurich
d) Catalan Institution for Research and Advanced Studies (ICREA) and Institute of Biotechnology and Biomedicine, Universitat Autònoma de Barcelona, ES-08193 Bellaterra (e-mail: xavier.daura@uab.es)

Under specific conditions, the complex formed by the adaptor protein Grb2 and the guanine-nucleotide exchange factor Sos2 is responsible for the activation of Ras, a low-molecular-weight GTPase involved in the control of cell proliferation and differentiation. The interaction between the N-terminal SH3 domain of the mouse Grb2 and one of its potential target sequences in the mouse, Sos2, a 15-residue peptide corresponding to residues 1264–1278, had been studied by NMR. However, the resulting data provided very limited information on the structure of the peptide and its interaction with the protein. Here, we present results from a molecular-dynamics simulation aimed at producing a realistic, atomic model for the interaction between the N-terminal SH3 domain of Grb2 and the SPLLPKLPPKTYKRE peptide from Sos2. In the simulation, the peptide adopts an extended conformation over the protein’s binding surface. The proposed polyproline-type-II helicity appears only locally, and the peptide displays substantial flexibility. It is found that the peptide residues Lys10 to Tyr12 could be responsible for most of the specificity of the interaction.

Introduction. – The small (ca. 60-residue), noncatalytic protein domains SH3 (Src-homology-3) mediate protein/protein interactions in the cytoplasm [1–5]. They have no fixed topological position in the protein and can fold into a compact and functional module independently of surrounding sequences. The function of SH3 domains is strongly related to their ability to bind Pro-rich motifs in target proteins [6–12]. SH Domains were first discovered in cytoplasmic protein tyrosine kinases such as the src oncogene product [13–16]. SH3 Domains are present in proteins of single-cell organisms such as yeast and amoeba, as well as in proteins of invertebrates and vertebrates. This indicates that the domain has a long evolutionary history and is involved in vital cellular processes. In higher eukaryotes, many SH3-containing proteins are involved in signal-transduction pathways [17][18]. One type of signalling proteins that has attracted particular interest are the so-called adaptor proteins [19][20]. They contain SH2 (Src-homology-2) and SH3 domains and exhibit no enzymatic activity. The growth factor receptor-bound protein 2 (Grb2) is one of the best studied adaptor proteins [21–24].

The Grb2 and mSos2 Proteins. Grb2 is a 25 kDa protein that links receptor tyrosine kinases to Ras signalling [23][25–29]. It is the mammalian homolog of Sem-5 from the nematode C. elegans [30] and of Drk from Drosophila [31][32]. Grb2 is a modular
protein composed of a central SH2 domain flanked by two SH3 domains. The linker regions that connect the two types of modules are flexible, and confer an overall flexibility to the protein that might be crucial for its biological function [33][34]. The SH3 domains of Grb2 bind, among other proteins, the guanine-nucleotide exchange factors mSos1 and mSos2 [12][35 – 39], two mammalian homologs of ‘Drosophila’s Son of sevenless’ (Sos) [40] that are responsible for the activation of Ras. The binding is bivalent, i.e., the N- and C-terminal SH3 domains bind to noncontiguous Pro-rich regions on the target [41][42]. Grb2 and the Mammalian variants of Sos exist as preformed complexes in the cytoplasm of unstimulated cells [38]. In the presence of the appropriate growth factor, growth-factor receptors (receptor tyrosine kinases) are activated by a ligand-induced autodimerization mechanism [43][44]. Upon activation, receptor tyrosine kinases undergo rapid autophosphorylation on numerous, specific tyrosine residues. The autophosphorylation sites located outside the catalytic domain are usually involved in the recruitment of cellular target proteins. Thus, the SH2 domain of Grb2 binds to pYxNx motifs (where pY is a phosphorylated Tyr residue and the x and x residues any amino acid) on the activated receptors and on other specific proteins that are phosphorylated upon receptor activation [45 – 50]. In either case, the Grb2/Sos complex is recruited to the plasma membrane, into proximity with (membrane-bound) Ras [12][51]. The interaction between Sos and Ras promotes the conversion of Ras from its inactive GDP-bound form to its active GTP-bound form [18][22][25]. In turn, Ras-GTP activates a highly conserved kinase cascade that relays information from the cell surface to a variety of cytoplasmic and nuclear targets, controlling cell proliferation and differentiation [27][52][53].

The SH3 Fold. The first reported structures of an SH3 domain were from Src [54] and spectrin [55]. The structures of several other members of the SH3 family were shortly after determined [6][56 – 64]. Despite the low degree of sequence similarity, the basic fold of SH3 domains is conserved, with some variations in secondary structure length and composition. It consists of an antiparallel five-stranded β-sandwich, where the two β-sheets pack against each other at approximately a right angle [4][65][66]. Based on the first descriptions of the SH3 structure, the strands are termed βa (at the N-terminus) to βe (at the C-terminus) [63] (Fig. 1). One of the β-sheets, βI, is formed by strands βa, βb1 (first half of βb), and βc, while the other, βII, is formed by βb2 (second half of βb), βd, and βd. This arrangement brings the junctions of the domain close together. Additional features of the canonical SH3 structure are the RT (Arg-Thr) loop, connecting strands βa and βb, the n-Src (neuronal-Src) insertion point between strands βb and βc [67], and a short segment of a 310-helix that separates the βd and βe strands. The RT loop resembles an irregular β-hairpin structure and has a variable sequence among SH3 domains. A consequence of this particular fold is that a majority of the conserved residues in the SH3 domain are positioned close to each other [68][69], resulting in the formation of a hydrophobic patch on the protein surface, opposite to the N- and C-terminals, which constitutes the ligand-binding site.

SH3 Ligand Binding. Isolated SH3 domains bind Pro-rich peptide fragments with relatively low binding affinity (K_D ca. 1 – 100 μM) and selectivity [3][8][65]. Early studies identified the consensus sequence ΦPxΦP (where Φ is a hydrophobic residue, often Leu) as a core-conserved binding motif in SH3-binding peptides [11][70]. More recently, some sequences lacking this motif have been shown to also bind SH3 domains.
Determination by nuclear magnetic resonance (NMR) [72] and X-ray crystallography [73] of the first structures of SH3 domains complexed with Pro-rich peptides showed that the peptides adopt a left-handed polyproline-type-II (PPII) helix, when bound to the SH3 surface, as had been proposed earlier [74]. The PPII helix is an extended, relatively rigid secondary structure with three residues per turn and no internal H-bonds [75]. It is roughly triangular in cross-section, with residues at positions $i$ and $i+3$ lying on the same ‘edge’ of the helix. In the complexes, one base of this triangle accommodates on the relatively flat ligand-binding surface of the SH3 domain. This consists of three shallow pockets or grooves defined by conserved aromatic residues. Two of the three ligand-binding pockets of the SH3 domain are occupied by the two $\Phi\Phi$ ‘dipeptides’ of the ligand’s core binding motif, which are in register on two adjacent turns of the helix. The third pocket (specificity pocket), flanked by the RT and n-Src loops, interacts in most cases with a basic residue in the ligand, distal to the $\Phi\Phi\Phi$ core. Investigations into the structure of other SH3/peptide complexes [76–80] showed that peptide ligands can bind in two reverse orientations with respect to the SH3 domain. In the $[+]$-orientation [78], the N-terminus of the ligand lies in the vicinity of the RT and n-Src loops, whereas in the $[-]$-orientation this position is occupied by the C-terminus. The orientation of the peptide is dictated by the location of the distal basic residue relative to the $\Phi\Phi\Phi$ core [76]. Thus, SH3-binding peptides are classified as class-I and class-II ligands, depending on their consensus binding sequence [72][81]: peptides with the motif $\Phi\Phi\Phi\Phi\Phi\Phi$ (where ‘+’ is a basic residue, often Arg) bind to SH3 in the $[+]$-orientation and correspond to class-I ligands, whereas peptides with the motif $\Phi\Phi\Phi\Phi\Phi\Phi$ bind in the $[-]$-orientation and correspond to class-II ligands. Because the PPII helix of an SH3 ligand is pseudosymmetric [78], some SH3 domains, like that from Src [76], have the capacity to bind peptides of either class, but the biological implications are unclear. Recent efforts have addressed how selectivity can be achieved both in nature and in vitro [3]. Intuitively, the core $\Phi\Phi\Phi\Phi\Phi$ scaffold restricts the variability that can be used to generate specificity. One obvious, but limited, source of specificity is the specificity

Fig. 1. Representation of the SH3 fold. The structure corresponds to the Grb2 N-SH3 domain (PDB entry 3gbq).
pocket of the domain [82][83]. Other ligand residues outside the core binding motif can also contribute specificity by interacting with residues on the SH3 surface outside the PPII-binding groove, such as the highly variable RT and n-Src loops [84][85]. Effective selectivity could be increased in vivo by a variety of means, such as compartmentalization of potential interaction partners, additive effects of multiple separate interactions between two partners, and the cooperative assembly of multi-protein complexes; alternatively, selectivity could be not strictly important in vivo, and regulation may be achieved by global changes in complex dynamic binding equilibria [3][86].

Solution Structure of the Grb2 N-SH3/mSos2 Complex. The structure of the complex between the N-terminal SH3 domain of the mouse Grb2 [87] and a 15-residue peptide (SPLPKLPPKTYKRE), corresponding to residues 1264 ± 1278 of the mouse Sos2 [35], was determined by Wittekind et al. [80] by means of NMR spectroscopy. A final ensemble of 29 structures was obtained, with an average atom-positional root-mean-square deviation (RMSD) from the average structure of 0.06 nm for the backbone atoms of the SH3 residues 1 – 54. The SH3 residues 27 – 35 (n-Src loop) show relative disorder. This is a region of the domain that appears to be inherently less well-defined than the rest, as has been observed with other solution structures [6][54][60][62][64][72][76]. Additionally, no NOE cross-peaks involving residues 55 – 59 at the C-terminus were observed. For the mSos2 peptide (which we shall call Sos2p), only NOEs involving the resonances of the three Leu residues could be assigned. Intermolecular NOE cross-peaks involving the Me H-atoms of Leu p4 (where the ‘p’ superscript is indicative for Sos2p) and Leu p7 were observed, whereas intermolecular cross-peaks involving the Me H-atoms of Leu p5 were not. As expected (the LPKLPPK core sequence identifies the Sos2p as a class-II ligand), a [-] orientation of the peptide on the SH3 surface produced the best agreement with the NOE distance restraints derived for Leu p4 and Leu p7. Since there were no NOE distance restraints for SH3 residues 55 – 59, and for Sos2p residues 1 – 3 and 8 – 15 (residues Leu p4 to Pro p8 constitute the core binding motif), their conformations were randomly distributed among the 29 structures. Other than some variation in the χ1 angle of Phe p9, the SH3 side chains with NOEs to the peptide (Tyr p7, Phe p9, Trp p6, and Tyr p5) are well-defined. The Tyr p7 and Tyr p5 aromatic side chains form a shallow binding subsite (S1) for the Leu p4 Me groups, while the Phe p9, Pro p8 and Tyr p5 side chains form a binding subsite (S2), with Pro p8 at its floor, for Leu p7. Although no structural information was obtained for Lys p10, mutagenic studies showed that a basic amino acid residue is required at this position to retain full binding potential [80]. If a left-handed PPII-helical backbone conformation is assumed for residues 2 – 11 of Sos2p, the side chain of Lys p10 can be easily positioned close to the specificity subsite (S3), encompassed by the highly (negatively) charged sequences Asp p14-Glu p16 and Glu p16-Asp p13 [80].

Molecular-Dynamics (MD) Simulation of the Grb2 N-SH3/mSos2 Complex. The NMR data obtained by Wittekind and co-workers [80] did not provide enough resolution to determine whether Sos2p adopts a PPII left-handed helical conformation upon binding to SH3, as observed for other Pro-rich peptides [72][73][88]. In addition, the role of Lys p10 in Sos2p binding to SH3 could not be determined. The MD simulation study presented here had two purposes: 1) to show that it is possible to model the structure and dynamic behavior of a protein/peptide complex from a crude knowledge
of the relative structure of the peptide (in this case, NMR data providing the orientation of the peptide and the positioning of just two peptide residues); 2) to answer the two major questions left open by the NMR study, i.e., the structure of the bound peptide and the binding mode to the S3 subsite. To our knowledge, this complex has never been studied before by MD simulation.

**Results and Discussion.** – A 10-ns MD simulation of the SH3/Sos2p complex in aqueous solution was performed with the GROMOS96 package of programs and the GROMOS-43A1 force field [89][90]. The simulation was started from the coordinates of model 1 in the Protein Data Bank entry 1gbr (Fig. 2). In this model, the only interactions between the Sos2p and SH3 molecules are those of Leu$^{p4}$ with Tyr$^7$ and Tyr$^{p5}$ in the S1-binding subsite, and Leu$^{p7}$ with Phe$^9$, Trp$^{p6}$, and Tyr$^{p3}$ in the S2 subsite.

**Overall Structure.** After 0.8 ns, the peptide already adopts an extended conformation over the surface of the protein (Fig. 2), similar to that adopted by the tighter-binding peptide mSos1(1135–1144) in complexes solved at higher experimental resolution [77][79][88]. The time evolution of the interaction energy between SH3 and Sos2p and the atom-positional root-mean-square deviation (RMSD) of the SH3 backbone from its initial model structure are shown in Fig. 3. Both quantities reach an equilibrium roughly after 1 ns from the start, although substantial fluctuations are apparent along the entire simulation. On the basis of Figs. 2 and 3, the period between 0.8 and 10 ns was chosen for the calculation of ensemble averages.

In Fig. 4, the violations of experimentally derived inter-H-atom upper-bound distances by the corresponding simulation ensemble averages are shown. From the 285 average distances evaluated from simulation (corresponding to H-atom pairs at a minimum residue-sequence distance of $i \cdots i + 2$), 217 satisfy the experimental upper bounds, 48 exceed the upper bound by less than 0.1 nm, twelve exceed the upper bound by less than 0.2 nm (and more than 0.1 nm), seven exceed the upper bound by less than 0.3 nm (and more than 0.2 nm), and one exceeds the upper bound by more than 0.3 nm. Of the eleven experimental intermolecular upper-bound distances, only one (Trp$^{p6}$·H$^{p2} \cdots$ Leu$^{p7}$·H$^{p4}$) is violated (by 0.06 nm). Clearly, the packing of the N-terminal $\beta$-strand ($\beta_4$; Fig. 1) in the simulation causes some average distances to be in conflict with the experimental data (Fig. 4). The interface between $\beta$-strands $\beta_b$ and $\beta_c$ is also a point with larger upper-bound-distance violations. Nevertheless, in all these regions, there are also H-atom pairs that are kept within the distance bounds in the simulation, indicating that the packing of the secondary structure elements is not being lost. In addition, Fig. 5 shows that the five $\beta$-strands of the protein are stable along the entire trajectory. On the contrary, the short 3$_{10}$ helix between $\beta$-strands $\beta_d$ and $\beta_e$ is only intermittently sampled. Note, however, that, according to the same criteria applied to the trajectory coordinates (using PROCHECK [91]), this helix is also missing in several of the NMR model structures, including model 1 (Fig. 4). The root-mean-square fluctuation (RMSF) of the positions of the C$^\alpha$-atoms is also plotted in Fig. 5 for the protein and the peptide. The $\beta$-strands are, as expected, regions of relatively low backbone fluctuations, while the highest flexibility appears to reside in the RT loop (turn region), in the n-Src loop, and in the turn connecting $\beta$-strands $\beta_c$ and $\beta_d$. The three residues of the peptide with lowest RMSF are Pro$^{p5}$, Lys$^{p10}$, and, more surprisingly, Tyr$^{p12}$. The low RMSF of Lys$^{p10}$ and Tyr$^{p12}$ are indicative of their active role in binding,
which will be further demonstrated by means of H-bonds and side-chain contacts (see Tables 1 and 2 below). The relatively large RMSF of Leu7, which, nevertheless, keeps its position in the S2 binding subsite (see below) is also worth mentioning.

Regarding the conformation of the bound peptide, the Ramachandran plots shown in Fig. 6 illustrate that, while the \( \phi/\psi \) angle pairs corresponding to the left-handed polyproline-type-II helix are sampled by all residues (1 and 15 excluded), the only residues that consistently stay in this conformation are the Pro moieties. Thus, it appears unlikely that at any time the peptide, or even a central part of the peptide, could be adopting an ideal polyproline-type-II helix.

**SH3/Sos2p Interactions.** The interactions between the two molecules are summarized in Tables 1 and 2. Table 1 reports the occurrence of intermolecular H-bonds,
defined with a maximum H–acceptor distance of 0.25 nm, and a minimum donor–H–acceptor angle of 135°. Table 2 reports the occurrence of intermolecular side-chain contacts with a maximum distance of 0.5 nm between any two side-chain atoms.

A number of things stand out from Tables 1 and 2. First, the interaction between Sos2p and SH3 takes place mainly at the side chain/side chain and side chain/backbone levels, i.e., there are no backbone/backbone H-bonds. The residues Tyr, Arg, and Asp in the RT loop (distal to the turn; Fig. 1) appear to play a central role in binding the N-terminal segment of the peptide (residues Ser to Pro), although the interaction is clearly nonspecific (polar side chains in the protein, mostly nonpolar in the peptide). Leu interacts with Tyr, which, together with Tyr, constitutes the S1 binding subsite. Asn and Tyr in the β-helix region are also in contact with the segment Pro to Pro. In addition, Leu is in contact with Phe (RT loop), Trp (N-terminus of βc), and Pro (C-terminus of βd), which define, together with Tyr, the S2 binding subsite. Other SH3 residues in contact with the central segment of the peptide (Leu to Pro) are
Thr12 and Glu16 in the RT loop (turn region), and Asn35 in the n-Src loop (or N-terminus of βc). As proposed previously [80], Lysp10 occupies the S3 or specificity subsite in the protein, defined by residues Asp14 to Glu16 in the RT loop (turn region) and Glu30 to

![Diagram](image1)

**Fig. 4. Violations of NMR-derived distances in the time interval 0.8–10 ns.** Some 68 violations were found from a total of 285 interatomic distances evaluated (between residues at a minimum sequence distance of i–i + 2). The dashed lines mark the end of the SH3 molecule (atoms 1 to 621) and the beginning of the Sos2p molecule (atoms 622 to 780). Black: violation < 0.1 nm; red: 0.1–0.2 nm; green: 0.2–0.3 nm; blue: > 0.3 nm. Upper and right-hand panels: secondary-structure map of the NMR structure (model 1 in PDB structure 1gbr; for color coding, see Fig. 5).

<table>
<thead>
<tr>
<th>Donor</th>
<th>Acceptor</th>
<th>Occurrence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trp³⁶</td>
<td>N⁴</td>
<td>Pro⁵⁰</td>
</tr>
<tr>
<td>Tyr³⁷</td>
<td>O²</td>
<td>Pro⁵⁰</td>
</tr>
<tr>
<td>Lys³⁰¹</td>
<td>N</td>
<td>Glu¹⁶</td>
</tr>
<tr>
<td>Lys³⁰¹</td>
<td>N</td>
<td>Glu¹⁶</td>
</tr>
<tr>
<td>Thr³¹¹</td>
<td>O²</td>
<td>Asp²³</td>
</tr>
<tr>
<td>Thr³¹¹</td>
<td>O²</td>
<td>Asp²³</td>
</tr>
<tr>
<td>Tyr³¹²</td>
<td>O²</td>
<td>Asp²³</td>
</tr>
<tr>
<td>Tyr³¹²</td>
<td>O²</td>
<td>Asp²³</td>
</tr>
</tbody>
</table>

*) Only H-bonds present in at least 10% of the structures analyzed are reported.
Asp\textsuperscript{33} in the n-Src loop. In addition, Lys\textsuperscript{p10} is in contact with Trp\textsuperscript{36} (S2) and, via a prototypical cation/π interaction, with Phe\textsuperscript{47} (C98d). This type of interaction is rather remarkable, since it is not clear to what extend the type of force field used should be able to reproduce it. What was not predicted is that the rest of the C-terminal segment of the peptide also establishes specific contacts with the S3 subsite, especially Thr\textsuperscript{p11} with Asp\textsuperscript{33} and Tyr\textsuperscript{p12} with Asp\textsuperscript{15}. As a summary, the N-terminal segment of the peptide binds to the SH3 moiety mostly via nonspecific peptide backbone/protein side chain interactions, while the C-terminal segment binds mostly via specific peptide side chain/protein side chain interactions.

Conclusions. – The purpose of this study was to produce a realistic, atomic-resolution model for the interaction between the N-terminal SH3 domain of the adaptor protein Grb2 and the peptide SPLLPPKPPKTYKRE, corresponding to residues 1264–1278 of one of the proteins that bind to Grb2 in vivo, the guanine-nucleotide exchange factor Sos2. For a system involving a very flexible molecule (the peptide in this case), a realistic model can only be obtained on the basis of thermodynamic ensemble averages. The results reported here demonstrate that it is possible to model the structure and dynamic behavior of the SH3/Sos2p complex by
MD simulation from just a crude knowledge of the relative structure of the peptide, i.e., overall orientation with respect to the protein moiety and localization of two residues, Leu\textsuperscript{4} and Leu\textsuperscript{7}.

The NMR study conducted by Wittekind et al. [80] on the SH3/Sos2p complex left two major questions open due to the poor resolution obtained for the peptide: the overall structure of the peptide and the nature of the interactions with the specificity subsite (S3 binding subsite) of the protein. The ensemble of structures analyzed here, corresponding to a 10-ns trajectory of the SH3/Sos2p complex in aqueous solution, is not only in good agreement with the available experimental data, but, in addition, provides an answer to these questions. Thus, the peptide adopts an extended...
conformation over the surface of the protein, such that all peptide residues (with the exception of Glu51) are in direct contact with the protein. The previously proposed left-handed polyproline-type-II structure [80] is sampled by all peptide residues (1 and 15 excluded), but it is only predominant for the four Pro residues. As proposed by Wittekind et al. [80] on the basis of a hypothetical model, Lys p10 occupies the highly anionic S3 binding subsite. In addition, the simulation data reveal that Tyr p12 and, to a lower extent, Thr p11, bind also tightly to this site.

The authors thank Luciano Mueller for providing the X-Plor restraining files used in the derivation of the NMR model structures.

**Experimental Part**

**Molecular Modelling.** The simulation of the complex between the N-SH3 domain of Grb2 (SH3) and the mSos2 peptide (Sos2p) in explicit H2O was carried out with the GROMOS96 package of programs [89][90]. The molecular model belongs to the GROMOS96-43A1 force field [89]. In this force field, the aliphatic H-atoms are treated as united atoms together with the C-atom to which they are attached [92]. The initial coordinates for the SH3/Sos2p complex were taken from the Brookhaven Protein Data Bank (PDB), entry 1gbr, model 1. Residues – 8 A to 0 A, and 60 A to 65 A in 1gbr, corresponding to the cloning vector [80], were discarded in this study. The coordinates of polar H-atoms (bound to N- or O-atoms) and aromatic H-atoms of SH3 and Sos2p were generated from standard geometries. The tautomer NδH was chosen for His51 of SH3 [80]. The charge state of the ionizable groups was chosen in accordance to a pH of 6 [80].
The SH3/Sos2p complex was placed at the center of a periodic truncated octahedron. The dimensions of the box were chosen such that the minimum distance from any protein atom to the walls was 1.8 nm. The solvent was introduced into the box by using a (cubic) pre-equilibrated configuration of 216 SPC H2O molecules as a building block [95]. All H2O molecules with their O-atoms lying within 0.23 nm of a non-hydrogen protein atom were then removed. Truncated-octahedron periodic boundary conditions were applied from this point onwards.

A steepest-descent energy minimization of the system was performed to relax the first shells of H2O molecules around the protein. Protein atoms were positionally restrained by means of a harmonic interaction with a force constant of 2500 kJ mol⁻¹ nm⁻². Next, a steepest-descent energy minimization of the system without restraints was performed to eliminate any residual strain. The energy minimizations were terminated, when the energy change per step became smaller than 0.1 kJ mol⁻¹.

**Simulation Setup.** A 10-ns MD simulation of the SH3/Sos2p/H2O system at 303 K and 1 atm was performed. The initial velocities of the atoms were taken from a Maxwell–Boltzmann distribution at 100 K. The temp. of the system was kept at 100 K during the first 25 ps, and then raised to 200 K and to the final 303 K [80] in successive 25-ps intervals. During this initial 75 ps, all solute atoms were positionally restrained by means of a harmonic interaction with a force constant of 2500 kJ mol⁻¹ nm⁻², and relaxed to 250 kJ mol⁻¹ nm⁻² for the last 25 ps. The temp. and pressure were maintained to the desired values by means of temp. and pressure external baths [94]. The temp. of the solute and the solvent were, independently, weakly coupled to the temp. bath, with a relaxation time of 0.1 ps. The pressure of the system (calculated via a molecular virial) was weakly coupled to the pressure bath, with isotropic scaling and a relaxation time of 0.5 ps. An estimated value of 4.575 ¥ 10⁻² mol nm⁻³ was taken for the isothermal compressibility of the system at 303 K and 1 atm [89]. Bond lengths were constrained to ideal values [89] using the SHAKE algorithm [95], with a geometric tolerance of 10⁻⁴. A time step for the leap-frog integration scheme of 2 fs was used. The nonbonded interactions were evaluated by means of a twin-range method: the short-range Van der Waals and electrostatic interactions were evaluated at every time step by means of a charge-group pair list generated with a short-range cut-off radius of 0.8 nm. Longer-range Van der Waals and electrostatic interactions (between charge groups at a distance greater than the short-range cut-off, and smaller than a long-range cut-off of 1.4 nm) were evaluated every five time steps, at which point the pair list was also updated, and were kept unchanged between these updates. The cut-off radii were applied to the centers of geometry of the solute charge groups and to the O-atoms of the H2O molecules. Interactions beyond the long-range cutoff of 1.4 nm were approximated by a Poisson–Boltzmann reaction field [96–98], assuming an electrostatic continuum with the relative dielectric permittivity of the SPC H2O model (εr = 54.0) [99].

**Analysis.** Trajectory coordinates and energies were saved at regular intervals of 0.5 ps. Ensemble averages were calculated over the time period 0.8–10 ns. RMSD and Secondary-structure calculations as well as Ramachandran plots were based on structures extracted at 5- ps intervals. The rest of the analysis was based on structures extracted at 0.5- ps intervals. Least-squares fitting of atomic coordinates for the calculation of atom-positional RMSD and atom-positional RMSF was based on the backbone atoms (N, Ca, C) of all but the terminal residues of SH3 and Sos2p. H-bonds were calculated according to a geometric criterion: a H-bond was defined by a minimum donor–H–acceptor distance of 0.25 nm and a maximum H–acceptor distance of 0.25 nm.

Side-chain contacts were evaluated on the basis of a distance criterion: two side chains were considered to be in contact, if two (or more) of their respective atoms were at a distance shorter than 0.5 nm. Protein secondary structure was calculated with the program PROCHECK [91], with assignments according to Kabsch and Sander [100]. Inter-H-atom distances derived from the exper. NOE intensities [80] were compared to the corresponding average effective inter-H-atom distances in the simulation, calculated as <r²>⁻¹/². The original list of 886 exper. distance restraints was reduced to 285 after eliminating atom pairs involving residues 8 A to 0 A, and 60 A to 65 A, and atom pairs belonging to residues at a sequence distance shorter than i−i+2. As already mentioned, in the GROMOS96-43A1 force field, aliphatic H-atoms are treated as united atoms together with the C-atom to which they are attached. Inter-H-atom distances involving aliphatic H-atoms were, thus, calculated by defining virtual (for CH and prochiral CH2) and pseudo (for nonstereospecific CH2 and CH3) atomic positions for these H-atoms at the time of analysis. Pseudo atomic positions were also used in typically unresolved cases like, e.g., the C7 and C9 H-atoms of Val, or the C10 and C12 H-atoms of Phe [89].

**REFERENCES**


Received January 3, 2004