Role of Interfacial Water Molecules in Proline-rich Ligand Recognition by the Src Homology 3 Domain of Abl*

The interaction of Abl-Src homology 3 domain (SH3) with the high affinity peptide p41 is the most notable example of the inconsistency existing between the currently accepted description of SH3 complexes and their binding thermodynamic signature. We had previously hypothesized that the presence of interfacial water molecules is partially responsible for this thermodynamic behavior. We present here a thermodynamic, structural, and molecular dynamics simulation study of the interaction of p41 with Abl-SH3 and a set of mutants designed to alter the water-mediated interaction network. Our results provide a detailed description of the dynamic properties of the interfacial water molecules and a molecular interpretation of the thermodynamic effects elicited by the mutations in terms of the modulation of the water-mediated hydrogen bond network. In the light of these results, a new dual binding mechanism is proposed that provides a better description of proline-rich ligand recognition by Abl-SH3 and that has important implications for rational design.

The recognition of proline-rich sequences by protein-protein interaction modules, such as SH3 or WW domains, is one of the most common mechanisms by which specific, transient protein-protein interactions are established within the cell. SH3 domains are found in oncoproteins and proteins overexpressed in deregulated signaling pathways during cancer development and are also associated with other pathologies such as AIDS, osteoporosis, or inflammatory processes (1–2). Inhibitors of the interactions between SH3 domains and their partners have proved to be promising therapeutic agents, validating these domains as attractive targets for drug design (3–6). Nevertheless, despite the wealth of structural and functional information collected during the last 2 decades, the forces driving proline-rich ligand recognition by SH3 domains are still not fully understood.

SH3 domains fold into a β-barrel structure composed of two orthogonal anti-parallel three-stranded β-sheets connected by three main loops (RT, n-Src, and distal loops). The binding site is a relatively flat, hydrophobic surface that consists of three shallow pockets. SH3 ligands typically contain the φPφPφP motif (where φ and p are frequently hydrophobic and proline residues, respectively) and bind in a PPII conformation so that each of the φP moieties packs tightly into one hydrophobic pocket formed by highly conserved aromatic residues on the surface of the domain. Additional interactions, which have been proposed to confer increased affinity and specificity, are established between residues flanking the core motif in the ligand and a third pocket delimited by the RT and n-Src loops, whose sequences vary among different SH3 domains (7–9).

According to this description, the recognition of proline-rich sequences by SH3 domains, based primarily in the burial of hydrophobic surfaces in the ligand and SH3-binding site (7, 9), would be expected to present a thermodynamic signature dominated by the hydrophobic effect, with the main driving force being a favorable entropic contribution opposed by positive, unfavorable enthalpies of binding (10–11). Nonetheless, all thermodynamic studies of SH3 ligand binding reported to date have surprisingly revealed an invariantly negative binding enthalpy that is partially compensated by unfavorable entropic contributions (12–17). This thermodynamic behavior, which cannot be rationalized exclusively in terms of direct interactions between hydrophobic surfaces, reveals an underlying complexity in the recognition of proline-rich ligands by SH3 domains. Additional factors, such as the modulation of SH3 dynamics, the redistribution of the native state ensemble upon ligand binding, or the impact of the conformational equilibrium of the peptide ligand itself on the binding energetics, have been proposed to contribute to the observed thermodynamic behavior (13, 16).

The binding of the high affinity proline-rich peptide p41 (APSYSPPPPP) (18, 19) to the SH3 domain of Abl is the most striking example of the inconsistency between the currently accepted description of proline-rich ligand recognition by SH3 domains and their thermodynamic signature. Despite the
Interfacial Water Molecules in Abl-SH3 Complexes

highly hydrophobic character of the ligand, this interaction is characterized by an extremely negative binding enthalpy (−92 kJ·mol⁻¹) (12), which is notably bigger than the values reported for other SH3 complexes that typically range between −20 and −50 kJ·mol⁻¹ (13–17). In previous work, we identified a set of fully buried water molecules at the binding interface that mediate the interactions between the peptide ligand and the SH3 domain and postulated that these water-mediated hydrogen bonds, which generate an extended and considerably more polar binding interface, are key for understanding the thermodynamic behavior of the system (12). To test our hypothesis and get a better insight into the role of interfacial water on proline-rich ligand recognition, we have carried out a thorough thermodynamic, structural, and computational study of the interactions of p41 with Abl-SH3 and a set of mutants rationally designed to perturb the network of water-mediated hydrogen bonds. Our results confirm the relevance of water-mediated interactions in Abl-SH3 complexes, providing a rationalization for the thermodynamic effects and a detailed description of the properties of the different hydration sites of interest for rational ligand optimization. In the light of these results, a new paradigm for proline-rich ligand recognition by Abl-SH3 emerges, implying a dual binding mechanism that combines the canonical hydrophobic interactions with the establishment of a network of peripheral water-mediated hydrogen bonds. This dual binding mechanism has important implications for the development of new ligand design strategies and for the understanding and modulation of SH3 signaling networks.

EXPERIMENTAL PROCEDURES

Protein and Peptide Samples—Mutants of the Abl-SH3 domain were constructed using the QuikChange site-directed-mutagenesis kit (Stratagene) and the WT Abl-SH3 plasmid as template. Oligonucleotides were purchased from Bonsai Technologies (Madrid, Spain). WT and mutants of the Abl-SH3 domain were expressed and purified as described previously for the wild type protein (12). The peptide p41 (Ac-APSYSPPP-P-NH₂) was bought from Sigma Genosys with a purity of >95%. Protein concentration was determined by absorbance at 280 nm using an extinction coefficient of 16,900 M⁻¹·cm⁻¹. Peptide concentration was determined by absorbance at 278 nm using an extinction coefficient of 1450 M⁻¹·cm⁻¹. Extinction coefficients were calculated using the Gill and von Hippel's method (20).

Isothermal Titration Calorimetry—Isothermal titration calorimetry was performed using a high precision MCS titration calorimetric system (Microcal Inc., Northampton, MA) as described previously (12). The heat produced by the binding reaction between the Abl-SH3 domain and the peptide ligand was calculated as the difference between the heat of reaction and the corresponding heat of dilution, which were obtained from independent titrations of the peptide ligand into the buffer. The resulting binding isotherms were analyzed by nonlinear least square fittings of the experimental data to a model corresponding to a single set of identical sites, as described before (12).

Differential Scanning Calorimetry—Differential scanning calorimetry experiments were performed in a VP-DSC microcalorimeter (Microcal) at a scan rate of 90 K·h⁻¹. The protein concentration in all DSC experiments was ~3 mg·ml⁻¹. The temperature dependence of the molar partial heat capacity (C_p) of the SH3 domains was calculated from the DSC data and analyzed using Origin 6.1 (OriginLab). C_p curves were fitted by a nonlinear least squares method using the two-state unfolding model as described elsewhere (21).

Crystallization and Data Collection—Crystals of free Abl-SH3 mutants were obtained using the sitting-drop vapor-diffusion method at 15 °C by mixing an equal volume of protein at 10–20 mg·ml⁻¹ and precipitant solution. Crystals grew in a broad range of pH values (3–8) in 2 M ammonium sulfate, 5% PEG300, 10% glycerol, and 0.1 M buffer solution (sodium citrate and HEPEs in the pH range 3–5 and 6–8, respectively). The N114Q-p41 complex was obtained by mixing the protein, dialyzed against 50 mM glycine, pH 3.0, at 8 mg·ml⁻¹, with lyophilized p41 in a 1:2 molar ratio. Crystals were obtained in 2 M ammonium sulfate, 0.4 M NaCl, 0.1 M sodium citrate, 10% glycerol pH 3.5. X-ray diffraction data were collected, integrated, and scaled as described before (22), although in this case data collection was performed in a cold nitrogen stream at 100 K. A summary of the data collection statistics is shown in Table 1.

Structure Resolution and Refinement—Initial phasing was obtained using MOLREP (23) together with the coordinates for the WT Abl-SH3-p41 complex (PDB code 1bbz) (19), from which the peptide ligand and water molecules had been removed. The molecular replacement solution refinement was conducted using several cycles of restrained positional and temperature factor refinement in REFMAC5 from CCP4 suite (24) in alternation with manual building using the resulting αA-weighted (2F_o − F) and (F_o − F) electron density maps in the COOT program (25). Water molecules were placed in the electron density difference maps using the ARP/wARP version 5.0 program from the CCP4 suite. The quality of all structures was checked using PROCHECK (26). All residues were found in the most favored regions of the Ramachandran plot. Structure refinement statistics are compiled in Table 1.

Accessible Surface Area Calculations—Changes in accessible surface area were calculated according to the Lee and Richards' algorithm (27) as described before (12).

Molecular Dynamics Simulations—MD simulations were carried out with the AMBER 8.0 package using as initial configurations the x-ray structures for the WT-p41 (PDB code 1bbz), N114A-p41 (PDB code 2o88), and N114Q-p41 complexes. All crystallographic waters were removed except for those buried at the binding interface. All hydrogen atoms were added using the Xleap tool. Standard ff03 force field parameters were used. Each complex was solvated in a truncated octahedron periodic box filled with TIP3P water molecules and neutralized by counter ions. MD simulations were preceded by two energy-minimization steps as follows: 1000 cycles (500 steepest descent and 500 conjugate gradients) with 500 kcal·mol⁻¹ harmonic force restraints on ligand and protein atoms followed by 2500 cycles (1500 steepest descent and 1000 conjugate gradients) without constraints. The system was heated from 0 to 300 K for 20 ps and equilibrated for 30 ps. 12 ns of productive MD runs were carried out in periodic boundary conditions in an isothermal isobaric ensemble (NPT) at 1 atm, with 2-fs time integration.
Interfacial Water Molecules in Abl-SH3 Complexes

TABLE 1
X-ray data collection and refinement statistics

| N114Q-p41 complex | N114Q pH 3 | N114A pH 3 | N114A pH 7 | N114F pH 7 |
|-------------------|-----------|-----------|-----------|-----------|
| **Space group**   | P2₁2₁2₁  | C222      | C222      | P3₁₂      | P3₁₂      |
| **Cell dimensions** |           |           |           |           |
| a, b, c           | 45.99, 47.64, 55.66 Å | 44.68, 53.00, 41.10 Å | 41.30, 44.82, 53.10 Å | 51.35, 51.35, 46.39 Å | 49.75, 49.75, 45.13 Å |
| α, β, γ           | 90, 90, 90° | 90, 90, 90° | 90, 90, 90° | 90, 90, 120° | 90, 90, 120° |
| **Resolution range** | 50.0 to 1.85 Å | 41.10 to 1.80 Å | 41.00 to 1.35 Å | 44.47 to 2.15 Å | 43.10 to 2.25 Å |
| Rmerge (%)        | 6.5% (24.3%) | 3.3% (26.9%) | 1.8% (12.8%) | 5.3% (30.0%) | 3.57% (28.35%) |
| I/σ(I)            | 19.2 (3.5) | 26.1 (3.9) | 32.8 (7.3) | 27.50 (3.97) | 22.75 (3.48) |
| Data completeness | 92.8% (72.5%) | 91.3% (67.9%) | 92.3% (82.2%) | 99.9% (100%) | 99.8% (98.4%) |
| Redundancy        | 7.60 (1.26) | 6.13 (1.11) | 6.79 (3.00) | 8.85 (4.07) | 9.38 (4.22) |

**Refinement**

|                  |                  |                  |                  |                  |
|-------------------|-------------------|-------------------|-------------------|-------------------|
| **No. of reflections** | 18 to 1.8 | 13.2 to 1.8 | 13.8 to 1.4 | 13.6 to 2.2 |
| Rmerge/Rfree (%)  | 19.2% (42.8%)    | 20.8% (22.8%)    | 20.8% (23.5%)    | 22.2% (25.4%)    |
|                  | 20.2% (28.6%)²   | 26.7% (30.6%)    | 23.9% (34.2%)    | 6.7% (34.6%)     |
|                  |                  |                  |                  | 28.3% (44.6%)    |

**No. of residues**

| Protein | 58 (chain A), 56 (chain B) | 63 | 63 | 56 | 56 |
|---------|--------------------------|---|---|---|---|
| Ligand/ion | 2                     | 1 | 1 | 2 | 1 |
| Solvent | 56                      | 23 | 36 | 19 | 17 |

**B-factors**

| Protein | 15.7, 16.1, 15.1, 15.8 | 16.1 | 15.2 | 19.8 | 17.3 |
|---------|------------------------|------|------|------|------|
| Ligand/ion | 52.4                  | 38.4 | 36.1 | 26.7 | 34.2 |
| Solvent | 24.5                   | 23.4 | 22.0 | 26.2 | 28.4 |

**r.m.s.d.**

| Bonds | 0.012 Å | 0.011 Å | 0.014 Å | 0.012 Å | 0.010 Å |
|-------|---------|---------|---------|---------|---------|
| Angles | 2.106° | 2.061° | 2.097° | 2.102° | 1.818° |

* R<sub>merge</sub> = ∑<sub>i</sub> ∑<sub>j≠i</sub> |I<sub>o</sub><sup>i</sup> - |I<sub>o</sub><sup>j</sup>|/ ∑<sub>i</sub> |I<sub>o</sub><sup>i</sup>, where |I<sub>o</sub><sup>i</sub>| is the weighted average intensity for all observations i of reflection h.

* The values in parentheses are for the highest resolution bin.

RESULTS

Asn-114 Mutants Alter the p41 Binding Energetics without Affecting the Conformational Equilibrium of Abl-SH3—In previous work (12), we identified a set of five fully buried water molecules at the binding interface of the complex between Abl-SH3 and the p41 peptide (19), which are numbered 1–5 in Fig. 1A. These water molecules are implicated in a complex network of hydrogen bonds that mediate the interactions between the peptide ligand and a set of residues in the SH3 domain different from those constituting the hydrophobic pockets for PPII recognition. These interfacial water molecules seem to be distributed in two distinct hydration regions defined by different water-coordinating amino acids as follows: (a) waters 1–3 in Fig. 1 that mediate the interactions between residues Asn-94, His-95, Asn-96, and Glu-98 in the n-Src loop and the specificity region of the ligand, and (b) waters 4 and 5 that bridge the interactions between residues Asn-114 and Ser-113 in the 3<sub>10</sub> helix and the PPII region of p41. These residues, which are not implicated in any direct contacts with the ligand, define an extended interaction surface characterized by a considerably higher polar character than the canonical binding site (analysis of the solvent-accessible area, ASA, shows that the ΔASA<sub>pol</sub>/ΔASA<sub>norm</sub> ratio drops from 2.9, characteristic of a highly hydrophobic interaction, to 1.7) that is in better agreement with the strongly exothermic character of the interaction.

To probe the role of the different hydration sites in the recognition of p41 and their influence on the binding energetics, we generated a full set of conservative mutants of the Abl SH3 domain designed to perturb the water-mediated hydrogen bond network with the minimum impact on the structural and conformational properties of the domain. This set included the elimination of the water-coordinating side chains at positions Asn-94, Asn-96, and Asn-114 as well as the modification of chain length at one representative position in each hydration region (N94T, N94Q and N114T, N114Q in the n-Src and 3<sub>10</sub> helix regions, respectively).

The binding energetics of p41 to the Abl-SH3 mutants were measured by isothermal titration calorimetry. As illustrated in Fig. 2A, most substitutions induced significant changes in the

...
Interfacial Water Molecules in Abl-SH3 Complexes

![Diagram of Abl-SH3 complex with water molecules]

**FIGURE 1.** Water molecules at the Abl-SH3/p41 binding interface for WT (a), N114A (b), and N114Q (c). The structure of the Abl-SH3 domain is shown in a gray schematic. Residues defining the canonical binding site for polyproline recognition are shown as gray sticks. The structure of the p41 peptide is shown as cyan sticks. Fully buried water molecules at the binding interface are shown as green spheres (sites occupied by water molecules are labeled from 1 to 5). Peripheral water-coordinating residues in the 3_10 and n-Src regions are shown as purple and dark pink sticks, respectively. Water-mediated hydrogen bonds are depicted as dotted green lines.

**FIGURE 2.** Effects of mutations on the binding energetics (A) and conformational properties (B) of the Abl-SH3 domain. A, differences in p41 binding Gibbs energy (striped bars), enthalpy (black bars), and entropy (gray bars) induced by the mutations with respect to the WT values. In all cases, the number of binding sites is 1 with the exception of mutants N94A, N94T, and N94Q that showed anomalously low values (n = 0.8, 0.9, and 0.5, respectively), indicating the susceptibility of the n-Src region to the substitutions. B, dependence of the denaturation enthalpy on the denaturation temperature (Tm) for Abl-SH3 mutants. The values for the partial molar denaturation enthalpy at the Tm (ΔHm) obtained from the best fits of the DSC curves to the two-state equilibrium model are plotted against the corresponding Tm values. With the exception of mutants N94A and N94Q, which are clear outliers, the enthalpy values show a linear dependence on Tm, indicating that their thermal unfolding can be described by a common enthalpy function. From the linear regression analysis of the data, a value of 4.7 kJ·(mol·K)^{-1} is obtained for the unfolding partial molar heat capacity that is, within experimental error, in good agreement with previous values for WT Abl-SH3 (41). Together with the low number of binding sites obtained in the isothermal calorimetry titrations for these mutants (between 0.4 and 0.8 depending on the mutant), clearly indicates that the mutations introduced at the n-Src region induce a significant perturbation of the Abl-SH3 conformational equilibrium, which greatly complicates the interpretation of their effects on the binding energetics. Conversely, the substitutions at position Asn-114 in the 3_10 helix region were very well tolerated. These mutants presented thermodynamic parameters very similar to the WT values (ΔTm, ΔH_{Tm}, and ΔG_{25°C} values below 2 °C, 10 and 4 kJ·mol^{-1}) and could be described by a common enthalpy function, indicating that the mutation of the Asn-114 side chain does not significantly perturb the enthalpic contributions of intra-domain interactions (see Fig. 2B). Consequently, the set of mutants at position Asn-114 constitute an ideal instrument to probe the thermodynamic effects associated with the reorganization of the water-mediated interaction network.

In this sense, the elimination of the Asn-114 side chain in the N114A mutant leads to a considerable reduction in the favorable binding enthalpy (ΔH_{N114A} = 5.6 kJ·mol^{-1}), which is probably associated with the release of interfacial water molecules or at least the weakening of their interactions due to the loss of the two water-mediated hydrogen bonds observed between Asn-114 and water 5 in the crystal structure. On the other hand, slightly more favorable binding enthalpies were obtained for the N114T and N114Q mutants (ΔH_{N114T} = -3.2 kJ·mol^{-1} and ΔH_{N114Q} = -2.4...
Asn-114 Mutants Induce Very Subtle and Local Effects on the Abl-SH3 Structure—To further confirm the structural integrity of the N114A mutants and investigate the relationship between the observed thermodynamic effects and the loss or modulation of the interfacial network of water-mediated interactions, we solved the crystal structures of several Asn-114 mutants, both unbound (N114A, N114Q, and N114T) and in complex with p41 (N114A-p41 (22) and N114Q-p41), at high resolution. These structures confirm that, as indicated by the DSC results, the substitution of N114 by Ala, Gln, or Thr does not induce significant structural changes in the SH3 domain. The overall structure is maintained between the different mutants with Ca r.m.s.d. values with respect to the WT structures of 0.46–0.94 Å for the unbound proteins and 0.34–0.58 Å for the complex structures (see Fig. 3). For the p41 peptide, the SH3 residues defining the canonical hydrophobic binding site and the water-coordinating side chains adopt very similar conformations in the three complexes, as illustrated in Fig. 1, b and c.

Occluded water molecules were also found at equivalent positions in the three complexes (WT-p41, N114A-p41, and N114Q-p41), with the exception of water at site 3, characterized by a high B factor in the WT-p41 complex, which was absent from the mutant structures. As is typical for immobilized waters, the B factors for all other water molecules were low and similar to those of the residues in the peptide and the SH3 domain that coordinated them (Fig. 4). The identities, solvent accessibilities, and B factors for all the interfacial water molecules are summarized in Table 2. It is interesting to note that the presence of most of the water molecules at the binding interface seems to be contingent on the binding of p41, with the exception of waters at hydration site 4, coordinated by the backbone atoms of residues Asn/Ala/Gln-114, Ser-113, and Glu-98, which are observed in most structures of free Abl-SH3 variants (see Fig. 3c). Remarkably, the B factors of waters at position 4 in the free Abl-SH3 domains are very similar to those of the equivalent waters in the complexes, indicating a ligand-independent tight coordination.

The pattern of water-mediated interactions is mostly conserved in the WT, N114A, and N114Q complex structures, with the effects of the substitutions being exclusively restricted to the mutation site. Despite the loss of the two hydrogen bonds established between the Asn-114 side chain and water 5 in the WT-p41 complex, water 5 remains at the binding interface in the N114A-p41 structure coordinated only by two hydrogen bonds, one of which is established with the carbonyl oxygen of P7 in the p41 peptide and the other of which is established with the water molecule at site 4. As mentioned before, the loss of water-mediated hydrogen bonds correlates well with the less favorable binding enthalpy obtained for this mutant. Unfortunately, the structure of the N114Q-p41 complex is not very informative about the changes induced by the mutations on the water-mediated interactions, because the Gln-114 side chain is oriented toward the bulk solvent in a conformation stabilized by crystal contacts. Despite this, most WT water-mediated hydrogen bonds do not directly implicating Gln-114 are preserved in this structure.

Effects on the Binding Energetics Correlate with Changes in the Dynamic Properties of Waters at Site 5—The static picture provided by the crystal structures is not sufficient to fully understand a highly dynamic process such as the continuous exchange of water molecules between the bulk solvent and the binding interface. Thus, to further characterize the network of water-mediated interactions and gain a better insight into the properties of the different hydration sites, the structure and dynamics of the WT-p41, N114A-p41, and N114Q-p41 complexes were also investigated by MD simulations. In all MD trajectories, atomic positions remained close to the starting crystallographic structures, with average Ca r.m.s.d. fluctuations of 1.8, 1.5, and 1.4 Å for the WT-p41, N114A-p41, and
three complexes (WT-p41, N114A-p41, and N114Q-p41). Water entry in this pocket was assisted by the side chain of Glu-98 that changed conformation continuously throughout the simulation, driving water molecules from the bulk solvent toward this region of the binding site. Water molecules in this cluster were in a relatively fast exchange with the bulk solvent, so that a high number of water molecules occupied hydration sites 1 and 2 with average residence times below 200 ps (see Fig. 5). Additionally, other secondary hydration sites in this region were occupied with less frequency, so that at particular moments the n-Src cluster contained up to 5 water molecules. The most relevant of these secondary sites is position 3 in Fig. 1, which was occupied about 10% of the simulation time in all complexes. This is in good agreement with the high B factor value of the corresponding water molecule in the WT-p41 crystallographic structure and the fact that this molecule is not observed in the N114A-p41 and N114Q-p41 crystal structures. As summarized in Table 3 and Fig. 5, the occupancies, amplitudes of fluctuations, and residence times of all water molecules in the n-Src cluster are essentially unaffected by the mutations at position 114 in the 3_10 helix region, highlighting the independent behavior of the two hydration clusters.

The 3_10 helix cluster is composed of two hydration sites (corresponding to crystallographic waters 4 and 5 in Fig. 1) that present quite distinct properties. Hydration site 4 was occupied over 98% of the simulation time by a small number of long lived water molecules in the WT-p41, N114A-p41, and N114Q-p41 complexes. For example, in the WT-p41 simulation most of the occupancy was contributed by only four molecules that stayed bound at the ligand-protein interface for periods that ranged between 1.5 and 2.7 ns. Water molecules at this site are characterized by a very low motional flexibility, as reflected in the small r.m.s.d. values (0.71 Å) and the low MD-derived B factor (17 Å²), which is in good agreement with the crystallographic results. As illustrated in Figs. 6 and 7, waters at site 4 are mostly bound to the protein and are held at this position by hydrogen bonds established with the backbone atoms of Glu-98 and Ser-113 and, less frequently, with the side chain of Asn-114. This is in good agreement with the fact that this hydration site is occupied in most Abl-SH3 crystal structures, independently of the presence of the ligand. Taken together, these observations clearly indicate that water molecules at site 4 should be considered as structural waters that constitute an integral part of the Abl-SH3 domain.

The situation is very different for hydration site 5. In the WT-p41 trajectory, this site was occupied about 80% of the simulation time by water molecules that were in faster exchange with the bulk solvent and showed higher motional flexibility (calculated B factor value of 26 Å²). These molecules interacted mostly with the carbonyl oxygens of P7 (70%) and P6 (15%), and were thus bound to the ligand. In fact, the hydrogen bonds observed in the crystal structure with the Asn-114 side chain were rarely established (see Figs. 6 and 7). Also, it is important to stress that no significant interaction between the two water molecules at the 3_10 helix cluster was observed for the WT-p41 complex (the direct hydrogen bond between water molecules at position 4 and 5 was established less than 5% of the simulation time).
In addition to their distinct dynamic properties and interaction profiles in the WT-p41 complex, waters at hydration sites 4 and 5 also showed very different responses to the mutations introduced at position Asn-114. As happened with n-Src hydration cluster, the dynamic properties of waters at site 4 were mostly unaffected by the mutations. Conversely, significant changes in the behavior of water molecules at site 5 were observed in the N114A-p41 and N114Q-p41 complexes. As illustrated in Fig. 7, both mutations resulted in a significant reorganization of the water-mediated interaction network that correlated well with the observed thermodynamic effects. In the WT-p41 complex, the Asn-114 side chain is implicated in several hydrogen bonds as follows: a direct interaction with the carbonyl oxygen of P7 in the p41 peptide and three hydrogen bonds established with water molecules at sites 4 and 5.

Even though these interactions are not very frequent (between 5 and 30% of the simulation time), the elimination of the Asn-114 side chain in the N114A complex leads, in combination with an increment in solvent accessibility in the 3₁₀ helix pocket, to significant changes in the overall pattern of water-mediated hydrogen bonds. The time of interaction of waters at site 5 with the carbonyl oxygens of P6 and P7 in the peptide ligand is considerably reduced. These interactions, especially the W₅-P₇ hydrogen bonds, are characterized by close to optimal geometries in the WT-p41 complex and thus would be expected to contribute strongly to the binding enthalpy. In compensation, the hydrogen bond between waters 4 and 5, which was very rare in the WT-p41 complex, is observed with a much higher frequency (40% of the simulation time), illustrating the robustness and plasticity of the water-mediated hydrogen bond network. In summary, the weakening of the

---

### Table 2

Crystallographic $B$ factors and solvent accessibilities of the water molecules at the Abl-SH3/ligand interface

| Water # | WT Abl-SH3 | N114A | N114Q | Kinase |
|---------|------------|--------|--------|--------|
|         | (a/b)      | (c/d)  | (e/f)  | (g/h)  |
|         | (Å²)       | (Å²)   | (Å²)   | (Å²)   |
| 1       | w1015      | w1036  | w1019  | w2036  |
|         | 20.91      | 24.15  | 20.66  | 27.36  |
|         | (0.03 Å²)  | (1.12 Å²) | (3.76 Å²) | (2.85 Å²) |
| 2       | w1064      | w1025  | w1028  |
|         | 36.30      | 22.04  | 32.24  |
|         | (0.00 Å²)  | (0.00 Å²) | (0.00 Å²) |
| 3       | w2103      | w2016  | w1003  |
|         | 44.07      | 47.21  | 37.88  |
|         | (5.55 Å²)  | (16.73 Å²) | (8.93 Å²) |
| 4       | w1082      | w1067  | w1097  | w1001  |
|         | 23.12      | 15.90  | 22.20  | 14.46  |
|         | (2.37 Å²)  | (4.89 Å²) | (0.00 Å²) | (6.25 Å²) |
| 5       | w1105      | w1060  | w1061  | w1018  |
|         | 22.56      | 21.70  | 13.32  | 21.74  |
|         | (0.00 Å²)  | (0.00 Å²) | (0.00 Å²) | (0.54 Å²) |
| 6       | w1089      |
|         | 22.64      |
|         | (0.00 Å²)  |

### B-factor range for protein residues

- Min: 4.6 Max: 44.7
- Min: 5.3 Max: 41.4
- Min: 5.3 Max: 43.3
- Min: 5.4 Max: 45.7
- Min: 8.4 Max: 47.0
- Min: 8.9 Max: 48.1
- Min: 9.6 Max: 44.3
- Min: 8.1 Max: 45.4
- Min: 13.6 Max: 75.0

---

1 Four structures for the WT-p41 complex in the asymmetric unit (PDB code 1bbz) are shown.
2 Two structures for the N114A-p41 complex in the asymmetric unit (PDB code 2o88) are shown.
3 Two structures for the N114Q-p41 complex in the asymmetric unit (PDB code 3eg1) are shown.
4 Water molecules at the binding interface between SH3 domains and the SH2 kinase linker in the crystal structure of the full-length Abl tyrosine kinase (PDB code 1opk) are shown.
5 The $B$ factor ranges correspond to the maximum and minimum values observed for SH3 residues for the structures refined without TLS.
strong interactions between waters at site 5 and the peptide ligand results in a 25% decrease in the water occupancy at this position and a less favorable binding enthalpy.

The N114Q substitution also induced a substantial reorganization of the interfacial hydrogen bond pattern. Nonetheless, in this case, the longer side chain, which, contrary to what was observed in the crystal structure, is mostly directed toward the binding pocket in the MD simulation, permits the establishment of a bigger number of highly geometrically optimized hydrogen bonds, resulting in a 15% increment in water occupancy with respect to the WT-p41 complex. As observed with the N114A-p41 complex, the N114Q mutation resulted in a decrease in the frequency of interaction of waters at site 5 with P7 in the peptide that was compensated by a striking (from 5 to 60%) increment in the time of interaction with waters at site 4 (see Figs. 6 and 7). Additionally, waters at site 5 established improved hydrogen bonds, as shown by their better distances and angles, with P6 in the peptide ligand and with the Gln-114 side chain itself with higher frequency. These optimized interactions justify the more favorable binding enthalpy obtained for this complex. Nonetheless, the weakening of the interactions with P7 in the ligand resulted in a sizeable increase in the water mobility and frequency of exchange at hydration site 5, which resulted in a higher entropic penalty that counterbalanced the enthalpic benefits.

**DISCUSSION**

The robustness of the 310 helix region of Abl-SH3 has allowed us to probe the energetic impact of altering the water-mediated hydrogen bond network at the Abl-SH3/p41 binding interface. In this respect, substitutions at position Asn-114 in Abl-SH3 resulted in significant enthalpic effects that correlated very well with the modulation of the water-mediated interactions. Even though the enthalpy changes were modest, and were in the lower limit of the values associated to the displacement of water molecules in other systems (28, 30), it is important to stress that they only reflect very subtle perturbations in the interactions within one of the hydration clusters, mostly associated with changes in the dynamic properties of waters at hydration site 5. If the observed effects on water occupancy at this position are extrapolated to the 80% occupancy found for the WT-p41 complex, we can tentatively estimate the contribution to the binding enthalpy of water molecules at site 5 to be around $-20 \text{kJ}\cdot\text{mol}^{-1}$. These results indicate that, according to our hypothesis, the interfacial water molecules contribute very substantially to the strongly exothermic binding enthalpy characteristic of Abl-SH3 interactions and need to be

#### TABLE 3

| Site | Occupancy | No. of molecules | $t_{\text{max}}$ (ps) | r.m.s.d.* (Å) |
|------|-----------|-----------------|-----------------------|--------------|
| **Site 1** | | | | |
| WT | 89.4 | 92 | 340 | 1.8 ± 2.3 |
| N114A | 88.7 | 114 | 834 | 1.5 ± 1.7 |
| N114Q | 93.2 | 101 | 568 | 1.1 ± 1.7 |
| **Site 2** | | | | |
| WT | 97.2 | 59 | 1584 | 1.3 ± 1.0 |
| N114A | 94.1 | 71 | 1560 | 1.3 ± 0.8 |
| N114Q | 91.1 | 67 | 1422 | 1.2 ± 0.6 |
| **Site 3** | | | | |
| WT | 7.2 | 35 | 82 | 1.7 ± 1.3 |
| N114A | 9.4 | 32 | 148 | 1.9 ± 1.2 |
| N114Q | 16.0 | 44 | 256 | 1.5 ± 0.7 |
| **Site 4** | | | | |
| WT | 97.6 | 12 | 2712 | 0.7 ± 0.5 |
| N114A | 97.2 | 9 | 3500 | 0.8 ± 0.5 |
| N114Q | 99.8 | 12 | 3332 | 0.8 ± 0.4 |
| **Site 5** | | | | |
| WT | 79.8 | 46 | 1214 | 1.2 ± 0.7 |
| N114A | 54.3 | 88 | 386 | 1.7 ± 2.8 |
| N114Q | 93 | 168 | 382 | 1.5 ± 2.0 |

* r.m.s.d. values were calculated taking as a reference the position of the water molecule at the mid-time of its residence at a particular hydration site.

**FIGURE 5.** Residence time distribution for the five hydration sites at the binding interface in the WT, N114A, and N114Q Abl-SH3-p41 complexes. Plotted are the number of water molecules within the different residence time ranges observed for each hydration site in the WT and mutant complexes. As can be observed, the dynamic properties of water molecules within the n-Src cluster (positions 1–3) are not affected by the mutations. Conversely, the mutations at position Asn-114 do have a significant impact on the residence time distribution of waters at the 310 helix cluster, especially in the behavior of waters at position 5.
included in the description of the binding interface for a complete understanding of these complexes.

The key role played by interfacial waters in Abl-SH3 opens new perspectives for rational design. In this respect, one of the main challenges faced when considering water-mediated interactions for ligand optimization is establishing which crystallographic water molecules are relevant, distinguishing between structural and displaceable waters. Toward this end, several computational and empirical procedures have been developed that are still yielding irregular results (31–32). Although considerably more laborious, the comprehensive analysis presented here provides a detailed and consistent characterization of the structural conservation and dynamic properties of the water molecules at the different hydration sites in Abl-SH3 complexes, establishing a solid background for devising new rational design strategies in Abl-SH3.

Our results confirm the existence of four well defined and robust hydration sites in the Abl-SH3-p41 complex that are highly conserved and occupied by fully buried water molecules with low B factors in most structures of the different Abl-SH3 complexes available to date (19, 22, 33–34), including the complexes between SH3 domains and SH2 kinase linkers in the context of the full-length Abl kinase. The relevance of these hydration sites in solution is further supported by the high level of occupancy (around 90% in all cases) observed in the MD simulations. From a functional standpoint, it is interesting to consider that the distinct dynamic properties of the $3_{10}$ helix and n-Src hydration clusters are imposed by the different conformational properties of these regions (35–36). Water molecules in the $3_{10}$ cluster (hydration sites 4 and 5) mediate the interactions between the most structurally stable (35–36) and highly conserved region in the SH3 domains and the highly structured PPII region in the ligand containing the polyproline consensus motif for SH3 recognition (9) and are thus very tightly coordinated. Conversely, water molecules in the n-Src cluster (hydration sites 1–3) are coordinated by residues in one of the most flexible and sequence-variable regions in the SH3 domain and the less structured residues in the specificity region of peptide ligand. Consequently, interfacial water molecules in this cluster transiently occupy different hydration positions with short residence times.

Our structural and MD analysis clearly shows that water molecules at hydration site 4 are tightly bound to the Abl-SH3

FIGURE 6. Properties and frequency of occurrence of hydrogen bonds at the binding interface of the WT (black bars), N114A (gray bars), and N114Q (white bars) Abl-SH3-p41 complexes. A, hydrogen bonds mediated by water molecules at position 4; B, hydrogen bonds mediated by water molecules at position 5; and C, direct hydrogen bonds. Shown are the weighted average (see “Experimental Procedures”) values for the angle and distance parameters corresponding to all hydrogen bonds established at the Abl-SH3-p41 binding interface with frequencies higher than 5% of the simulation time. Error bars correspond to the standard deviation values. Angles are reported according to AMBER-ptraj standards that assign to a linear H-bond ($180^\circ$) a value of $0^\circ$. In this context, values of $30 – 40^\circ$ correspond to close to optimal H-bond angles.
domain independently of the presence of the ligand, and are characterized by very high residence times (over 1–2 ns) and markedly small B factors in both complexes and free domains. Consequently, these waters should be considered as an integral part of the Abl-SH3 domain and should be included in the description of the binding site in structure-based design and ligand docking studies. Moreover, from a structure-based rational design perspective, water molecules at this site are optimal for the engineering of new, direct hydrogen bonds with the ligand. Because of their tight coordination and high stability, the conformational entropy associated with the establishment of new interactions with these water molecules will most likely be small. This is a very advantageous situation, because any enthalpic benefits arising from these new hydrogen bonds will efficiently translate into a net increment in binding affinity (37).

Independently of their dynamic properties, the presence of all other water molecules (waters at the n-Src cluster and water 5 at the 310 cluster) is contingent upon formation of the complex. This is consistent with the fact that waters at position 5 were found to interact mostly with the ligand. Waters at site 5 are retained at the binding interface in the mutant complexes because of the marked increase in the frequency of interactions with waters at site 4, further highlighting the central role played by these structural waters in Abl-SH3 recognition. Because of their looser coordination, water molecules at position 5 could constitute good candidates to be targeted for displacement. The comparison of the different structures of full-length tyrosine kinases provides a good illustration of this potential. Water molecules at position 5 are observed in all crystallographic structures of full-length Abl available to date (PDB codes 1opk, 1opl, and 2fo0) and are coordinated by a short serine side chain in the SH2-kinase linker that binds to the SH3 domain. Nonetheless, water molecules at this position are displaced by a longer glutamine side chain that occupies this position in the closely related c-Src tyrosine kinase (PDB code 2ptk).

In summary, the Abl-SH3-p41 complex is a clear example of how ignoring the contributions of interfacial water molecules can lead to incorrect or at least incomplete descriptions of the binding interface (38–39). The results of the thermodynamic, structural, and MD analysis presented here clearly demonstrate that, even though traditionally the recognition of proline-rich sequences by SH3 domains has been presented as a paradigm for hydrophobic binding, the interaction of p41 with Abl-SH3 takes place via a more complex mechanism in which the hydrophobic interactions established at the canonical binding site are complemented by a plastic and robust network of water-mediated hydrogen bonds to peripheral residues (wet spots) (39–40). This dual binding mechanism is in better agreement with the thermodynamic behavior of the system and has important implications for molecular modeling and ligand design.

Acknowledgments—We thank Prof. J. M. García-Ruiz and Dr. J. A. Gavira at the Laboratorio de Estudios Cristalográficos (Consejo Superior de Investigaciones Científicas) for technical assistance with crystalization and data collection within the project “Factoría Española de Cristalización” Ingenio/Consolider 2010. We also thank Dr. S. A. Samsonov for technical assistance with the molecular dynamics simulations and J. M. Martin-Garcia for many helpful discussions.

REFERENCES
1. Kay, B. K., Williamson, M. P., and Sudol, M. (2000) FASEB J. 14, 231–241
2. Dalgarno, D. C., Botfield, M. C., and Rickles, R. J. (1997) Biopolymers 43, 383–400
3. Feller, S. M., and Lewitzky, M. (2006) Curr. Pharm. Des. 12, 529–548
4. Oneyama, C., Agatsuma, T., Kanda, Y., Nakano, H., Sharma, S. V., Nakano, S., Narazaki, F., and Tatsuta, K. (2003) Chem. Biol. 10, 443–451
5. Vidal, M., Giguex, V., and Garbay, C. (2001) Crit. Rev. Oncol. Hematol. 40, 175–186
6. Posern, G., Zheng, J., Knudsen, B. S., Kardinal, C., Müller, K. B., Voss, J., Shishido, T., Cowburn, D., Cheng, G., Wang, B., Kruh, G. D., Burrell, S. K., Jacobson, C. A., Lenz, D. M., Zamborelli, T. J., Adermann, K., Hanafusa, H., and Feller, S. M. (1998) Oncogene 16, 1903–1912
7. Musacchio, A. (2002) Adv. Protein Chem. 61, 211–268
Interfacial Water Molecules in Abl-SH3 Complexes

8. Mayer, B. J. (2001) J. Cell Sci. 114, 1253–1263
9. Ball, L. J., Kühne, R., Schneider-Mergener, J., and Oschkinat, H. (2005) Angew. Chem. Int. Ed. Engl. 44, 2852–2869
10. Velázquez Campoy, A., Luque, I., and Freire, E. (2001) Thermochimica Acta 380, 217–227
11. Velázquez Campoy, A., and Freire, E. (2005) Angew. Chem. Int. Ed. Engl. 44, 2852–2869
12. Palencia, A., Cobos, E. S., Mateo, P. L., Martínez, J. C., and Luque, I. (2004) J. Mol. Biol. 336, 527–537
13. Ferreon, J. C., and Hilscher, V. I. (2004) Biochemistry 43, 7787–7797
14. Arold, S., O’Brien, R., Franken, P., Strub, M. P., Hoh, F., Dumas, C., and Ladbury, J. E. (1998) Biochemistry 37, 14683–14691
15. Renzoni, D. A., Pugh, D. J., Siligardi, G., Das, P., Morton, C. J., Rossi, C., Waterfield, M. D., Campbell, I. D., and Ladbury, J. E. (1996) Biochemistry 35, 15646–15653
16. Wang, C., Pawley, N. H., and Nicholson, L. K. (2001) J. Mol. Biol. 313, 873–887
17. Wittekind, M., Mapelli, C., Farmer, B. T., Suen, K. L., Goldfarb, V., Tsao, J., Lavoie, T., Barbacid, M., Meyers, C. A., and Mueller, L. (1994) Biochemistry 33, 13531–13539
18. Pisabarro, M. T., and Serrano, L. (1996) Biochemistry 35, 10634–10640
19. Pisabarro, M. T., Serrano, L., and Wilmanns, M. (1998) J. Mol. Biol. 281, 513–521
20. Gill, S. C., and von Hippel, P. H. (1989) Anal. Biochem. 182, 319–326
21. Viguera, A. R., Martínez, J. C., Filimonov, V. V., Mateo, P. L., and Serrano, L. (1994) Biochemistry 33, 2142–2150
22. Cámara-Artigas, A., Palencia, A., Martínez, J. C., Luque, I., Gavira, J. A., and García-Ruiz, J. M. (2007) Acta Crystallogr. D Biol. Crystallogr. 63, 466–465
23. Vagin, A., and Tsyblyakov, A. (1997) J. Appl. Crystallogr. 30, 1022–1025
24. Bailey, J. (1994) Acta Crystallogr. D Biol. Crystallogr. 50, 760–763
25. Emsley, P., and Cowtan, K. (2004) Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132
26. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallogr. 26, 283–291
27. Lee, K. H., Xie, D., Freire, E., and Amzel, L. M. (1994) Proteins Struct. Funct. Genet. 20, 68–84
28. Sharrow, S. D., Edmonds, K. A., Goodman, M. A., Novotny, M. V., and Stone, M. J. (2005) Protein Sci. 14, 249–256
29. Humphrey, W., Dalke, A., and Schulten, K. (1996) J. Mol. Graphics 14, 33–38, 27–28
30. Li, Z., and Lazaridis, T. (2005) J. Phys. Chem. B 109, 662–670
31. Raymer, M. L., Sanschagrin, P. C., Punch, W. F., Venkataraman, S., Goodman, E. D., and Kuhn, L. A. (1997) J. Mol. Biol. 265, 445–464
32. García-Sosa, A. T., Mancera, R. L., and Dean, P. M. (2003) J. Mol. Model 9, 172–182
33. Musacchio, A., Saraste, M., and Wilmanns, M. (1994) Nat. Struct. Biol. 1, 546–551
34. Nagar, B., Hantschel, O., Young, M. A., Scheffzek, K., Veach, D., Bornmann, W., Clarkson, B., Superti-Furga, G., and Kuriyan, J. (2003) Cell 112, 859–871
35. Sadqi, M., Casares, S., Abril, M. A., López-Mayorga, O., Conejero-Lara, F., and Freire, E. (1999) Biochemistry 38, 8899–8906
36. Casares, S., Ab, E., Eshuis, H., Lopez-Mayorga, O., van Nuland, N. A., and Conejero-Lara, F. (2007) BMC Struct. Biol. 7, 22
37. Lafont, V., Armstrong, A. A., Ohtaka, H., Kiso, Y., Mario Amzel, L., and Freire, E. (2007) Chem. Biol. Drug Des. 69, 413–422
38. Luque, I., and Freire, E. (2002) Proteins 49, 181–190
39. Teyra, J., and Pisabarro, M. T. (2007) Proteins 67, 1087–1095
40. Teyra, J., Doms, A., Schroeder, M., and Pisabarro, M. T. (2006) BMC Bioinformatics 7, 104
41. Filimonov, V. V., Azuaga, A. I., Viguera, A. R., Serrano, L., and Mateo, P. L. (1999) Biophys. Chem. 77, 195–208