Published in final edited form as:

Protein Pept Lett. 2012 August ; 19(8): 832–838.

Water-Refined Solution Structure of the Human Grb7-SH2 Domain in Complex with the erbB2 Receptor Peptide pY1139

Sally C. Pias, Dennis L. Johnson, David E. Smith, and Barbara A. Lyons*

Department of Chemistry and Biochemistry, New Mexico State University, Las Cruces, New Mexico

Abstract

We report a refinement in implicit water of the previously published solution structure of the Grb7-SH2 domain bound to the erbB2 receptor peptide pY1139. Structure quality measures indicate substantial improvement, with residues in the most favored regions of the Ramachandran plot increasing by 14 % and with WHAT IF statistics (Vriend, G. *J. Mol. Graph.* , **1990** 8(1), 52–56) falling closer to expected values for well-refined structures.

Keywords

Molecular dynamics simulations; NMR refinement; SH2 domain; water refinement

INTRODUCTION

The Grb7 protein is an intracellular adaptor molecule that serves regulatory and scaffolding roles in a number of signaling pathways. It has been noted for its function as a regulator of cell migration through binding to focal adhesion kinase (FAK) [1]. Grb7 is co-overexpressed with the erbB2 growth factor receptor in 20–30 % of breast cancers [2]. Further, overexpression of Grb7 is used as an indicator of metastatic potential in tumor tissues [1, 3]. The full-length protein consists of 535 amino acids and includes a single C-terminal SH2 domain (Src homology 2), which spans approximately 118 residues. (For a more detailed description of the Grb7 domain topology, please refer to our previous publications [4–5].) The Grb7-SH2 domain mediates binding to signal transducers in the cell membrane and is the target of cancer therapeutics in active development [6–7]. Accurate structural models of the ligand-bound Grb7-SH2 domain are important to facilitate the design of effective inhibitors and to further basic knowledge of SH2 domain function.

The structure of the Grb7-SH2 domain has previously been studied using both nuclear magnetic resonance spectrometry (NMR) and X-ray crystallography. Our group published a solution structure of the Grb7-SH2 domain in complex with pY1139, a phosphorylated ten-amino-acid peptide representative of the activated erbB2 receptor [5]. More recent crystallographic studies have produced structures of the Grb7-SH2 domain alone [7] and in complex with a nonphosphorylated inhibitor [8]. Although *in vitro* analyses indicate that the ligand-bound Grb7-SH2 domain is predominantly monomeric [7, 9–10], the NMR complex remains the only experimentally determined structure of a monomeric Grb7-SH2 domain with a bound ligand. We present a new ensemble of the Grb7-SH2/pY1139 complex, refined

in implicit water from the previously published ensemble using molecular dynamics simulated annealing with the original NMR-derived distance and dihedral angle restraints.

The original solution structure was determined in vacuum, using CNS software [11] with a relatively limited set of experimentally derived restraints. Since that structure's publication in 2003, the benefits of accounting for a protein's aqueous solution environment have become increasingly [12–14]. Though computationally costly, water refinement has become more common. Incorporating solvent effects can be especially valuable when restraints derived from experiment are in limited supply. In such cases, force field accuracy becomes more important than in instances where restraints are abundant, and structural quality may be substantially improved through implicit water refinement [13]. We have gained access to a parallel computing cluster, allowing refinement of the Grb7-SH2/pY1139 solution structure in implicit solvent with a state-of-the-art [15], all-atom force field.

RESULTS AND DISCUSSION

The NMR solution structure of the SH2 domain of human Grb7 in complex with the phosphorylated erbB2 receptor peptide pY1139 was refined using molecular dynamics simulated annealing within an aqueous environment simulated by a generalized Born (GB) implicit solvent model [16]. The GB model performs well in molecular dynamics simulated annealing NMR refinement, providing improvements on vacuum refinement comparable to those obtained using explicit solvent [12].

Water Refinement Increases the Conformational Diversity of Solvent-exposed Regions

The ensemble resulting from our molecular dynamics simulated annealing refinement of the Grb7-SH2/pY1139 complex is presented in (Fig. 1). Noteworthy in the new ensemble is a conformationally diverse region on the SH2 domain N-terminus, spanning residues 1–18 (blue trace in (Fig. 1B)). Although disordered in the original ensemble calculated in vacuum, this segment folded back against the globular portion of the domain in all of the original NMR models. In contrast, the new models, refined with implicit water, show residues 1–12 stretching away from the domain into the solvent and assuming a variety of orientations. Nuclear relaxation data support the likelihood of structural mobility in this region, with higher than average T_2 relaxation times and lower than average ^{15}N heteronuclear NOE (nuclear Overhauser effect) intensities observed for residues 1–12 of the Grb7-SH2 domain with and without binding to the pY1139 ligand [10]. Moreover, crystallographic study by Porter *et al.* [7] of the ligand-free Grb7-SH2 domain found a lack of electron density for residues 1–7, while the remaining residues in the region (through residue 18) assumed an orientation away from the globular portion of the domain, rather than flanking it as in the original NMR ensemble. Several additional solvent-exposed segments show substantial backbone conformational diversity in the new ensemble, colored orange in (Fig. 1B). These occur in the C-terminal region of the Grb7-SH2 domain and include amino acids in the D'E loop (residues 74–78), in the EF loop (residues 83–86), in the α -B helix (residues 92–104), and in the BG loop (residues 107–110). Our nomenclature follows that established for SH2 domains by Eck *et al.* [17].

The New Ensemble Represents a Refinement of the Grb7-SH2/pY1139 Solution Structure

Table 1 provides a comparison of structure quality indicators for the new and previously published NMR ensembles. These indicators consistently demonstrate improved quality in the new ensemble. First, the violation statistics show improvement in the overall number of NOE-derived distance restraint violations, with all restraints now satisfied within a 0.4-Angstrom margin. While the root-mean-square deviation (RMSD) to the ensemble average coordinates is somewhat greater for the new ensemble (0.99 Angstroms for backbone

atoms), the increase may be attributed to expected conformational variation caused by solvent effects. Increased RMSD has been observed in other studies where refinement is carried out in an aqueous environment, including a recent report by Schirra *et al.* [22]. The Ramachandran plot statistics show improvement, with 69.8 % of new ensemble residues in the most favored regions, as compared with 55.6 % of residues in the original ensemble. Finally, WHAT IF [20] Z-scores and RMS Z-scores for the new ensemble fall closer to the expected values for well-refined structures. We conclude, therefore, that the new ensemble represents a refinement of the NMR solution structure for the Grb7-SH2/pY1139 complex.

Conformational Diversity in key Sidechains of the Grb7-SH2 Domain Dimerization Interface Suggests an Explanation for Monomerization Upon Ligand Binding

Prior investigations have indicated that the free Grb7-SH2 domain exists predominantly as a dimer [23] but that the domain converts to monomeric form upon ligand binding [7, 9–10]. Indeed, the 2007 crystallographic study of the free Grb7-SH2 domain found dimers associated through a hydrophobic interface primarily composed of phenylalanine residues F90 and F99 on each monomer [7]. (The same residues are numbered F502 and F511 in that study). (Fig. 2A) demonstrates that F90 and F99 pack against one another in the crystal structure, with the F90 sidechain oriented perpendicular to the F99 ring. In the pY1139 peptide-bound, monomeric domain, these two phenylalanine sidechains take on a variety of orientations and fail to pack against one another in any NMR model (Fig. 2B). In addition, the *alpha*-B helix shows a somewhat different rotation in all models of the new NMR ensemble, placing the backbone atoms of F99 closer to the domain core than they appear in the crystal structure. The altered rotation of the *alpha*-B helix in the Grb7-SH2/pY1139 complex, coupled with the apparent mobility of the F90 and F99 sidechains, offers a possible explanation for the conversion of the Grb7-SH2 domain to monomeric form when bound to the pY1139 peptide [10]. Specifically, we propose that ligand binding induces a conformational change involving the rotation of the *alpha*-B helix and consequently impedes the formation of a stable dimerization interface.

It should be noted that the F99 sidechain was restrained in the original NMR structure calculation and in the water refinement by six ^1H – ^1H NOE-derived distances involving the F99 *beta*-hydrogen atoms, with restraint maxima ranging from 5.3 to 6.8 Angstroms. These are relatively long restraint distances, corresponding to relatively low NOE peak intensities and implying that the F99 sidechain is dynamic. Eighteen additional NOE distances restrained F99 backbone hydrogen atoms, with maxima ranging from 3.8 to 5.8 Angstroms. We previously reported that the ^{15}N heteronuclear NOE value for F99 is lower than average for the domain and that the F99 backbone nuclei exhibit higher than average T_2 relaxation times in both the ligand-bound and the ligand-free domain [10]. This combination of lower than average ^{15}N heteronuclear NOE values and higher than average T_2 relaxation times points to increased internal motion for the F99 residue backbone.

On a separate note, the F90 residue was restrained by five NOE-derived distance restraints, with maxima between 4.3 and 6.8 Angstroms. Seventeen additional distance restraints, with maxima from 3.3 to 6.8 Angstroms, were applied to F90 backbone atoms. In our previous nuclear relaxation study, we found for the Grb7-SH2/pY1139 complex that the F90 value for T_2 is lower than the molecule average and that the ^{15}N heteronuclear NOE value is higher than average. These relaxation measurements are consistent with relatively low mobility for the backbone of F90 and with the relatively low backbone diversity of our refined ensemble in the vicinity of this residue.

The Refined NMR Solution Structure Provides Valuable Ligand Binding Information

A notable feature of the Grb7-SH2 domain, in contrast with nearly all known SH2 domains, is its affinity for ligands bearing a motif that excludes a hydrophobic residue in the third position following phosphorylated tyrosine [24–26]. SH2 domains generally bind ligands with a linear conformation, in which the third residue after phosphorylated tyrosine (pY+3) contacts a hydrophobic binding pocket on the domain's surface. On the other hand, the Grb7-SH2 domain shows a preference for ligands that lack a pY+3 hydrophobic residue and that bind in a *beta*-turn or cyclic conformation [5]. We attribute this distinct binding preference to the behavior of additional residues that lengthen the Grb7-SH2 domain EF loop by four amino acids *versus* SH2 domain EF loops outside the Grb7 family. In both the original and the water-refined NMR ensembles, the EF loop aspartate residues D84 and D85 obscure the canonical hydrophobic binding pocket and appear to hinder its interaction with hydrophobic moieties on potential binding partners (Fig. 3B). As such, the EF loop adopts an “extended” conformation in both the original and the refined Grb7-SH2/pY1139 complex, and this loop conformation seems to be an important determinant of binding specificity. In contrast, the EF loop assumes a “bent” conformation in both the unliganded and the inhibitor-bound crystal structures (2QMS and 3PQZ), corresponding with substantially greater surface exposure of hydrophobic sidechains in the pY+3 binding pocket region ((Fig. 3A), showing the unliganded structure only; note its close conformational similarity with the inhibitor-bound structure [8]). The existence of two EF loop conformations (extended and bent) in the Grb7-SH2 domain is supported by prior backbone nuclear relaxation measurements, which point to increased mobility in the EF loop region [10]. Given the apparent consistency of the extended EF loop conformation with the Grb7-SH2 domain binding selectivity, we recommend its consideration in inhibitor design.

METHODS

NMR-derived Restraints

The present refinement uses the previously published [5] distance and dihedral angle restraints for the Grb7-SH2/pY1139 complex. Distance restraints were derived from NOE peak intensities, while the backbone dihedral angle restraints (*phi* only) were determined through analysis of J-coupling data. Restraint lists used in the current study were downloaded in an appropriate format from the BiomagRes-Bank (BMRB) [28].

Molecular Dynamics Simulated Annealing

All computation was performed on a cluster of SuperMicro servers, with a total of sixteen 1.0 GHz Xeon 5335 quadcore CPUs and 128 GB of memory. The refinement was carried out with AMBER software (version 9) [29], using the well-validated [15] ff99SB force field [30], along with a modified [31]–[32] generalized Born (GB) implicit solvent model (GB model) [16] and no cutoff for nonbonded interactions. The phosphorylated tyrosine residue on the pY1139 peptide was treated as fully unprotonated, using AMBER force field parameters contributed by Homeyer *et al.* [33]. The salt concentration parameter was set at 150 mM to match the NMR solution conditions [5]. The ten models of the originally published NMR ensemble for the Grb7-SH2 domain in complex with the pY1139 peptide [5] served as starting structures. Each NMR model was energy minimized using the GB model for 2000 steps, then subjected to two rounds of 32 ps molecular dynamics simulated annealing with the GB model and with NMR-derived distance and dihedral angle restraints from 800 K to 0 K. Additional restraints determined from the original ensemble were applied to maintain chirality and *trans-omega* dihedral angles for non-proline residues. Restraint force constants were adjusted to optimize both the fit to NMR-derived restraints and the structure quality. Following molecular dynamics simulated annealing, the best

energy structure for each model was minimized for 2000 steps with the GB model and with the restraints described above.

The H++ server (<http://biophysics.cs.vt.edu/H++>) was used to generate a consensus set of predicted protonation states for ionizable residues in all models of the Grb7-SH2/pY1139 complex [34]. Calculations were made using pH 6.6 and 150 mM salinity as in the NMR data acquisition [5] and using the Poisson-Boltzmann method as implemented in H++. All histidine residues were predicted to be singly protonated, with the exception of doubly protonated H100. In addition, glutamate residues E47 and E74 were predicted to have protonated side chain carbonyl groups. These predictions were consistent across all NMR models and were used throughout the refinement.

Our refined ensemble for the Grb7-SH2/pY1139 complex has been deposited in the Protein Data Bank, under the identifier **2I4K** and the new record is intended to replace the previously published ensemble (1MW4). Availability of the refined coordinates will facilitate investigations of the Grb7-SH2 domain function, including therapeutic inhibitor docking.

Acknowledgments

This work was carried out using computational resources provided by the New Mexico State University Department of Computer Science (Bioinformatics Cluster). It was supported financially by National Science Foundation awards 420-40-50 (IGERT training grant), HRD-0420407 (CREST Center for Research Excellence), and 0937060 (grant to the Computing Research Association for the CIFellows Project), as well as National Institutes of Health grant P20RR016480 (New Mexico INBRE network).

ABBREVIATIONS

Grb	Growth factor receptor bound
SH2	Src homology 2
erbB2	Erythroblastic leukemia viral oncogene homolog 2 (also known as HER2 = Human epidermal growth factor receptor 2)
FAK	Focal adhesion kinase,
NMR	Nuclear magnetic resonance spectrometry
NOE	Nuclear Overhauser effect
RMSD	Root-mean-square deviation
RMS	Root-mean-square

REFERENCES

1. Shen TL, Guan JL. Grb7 in intracellular signaling and its role in cell regulation. *Front. Biosci.* 2004; 9:192–200. [PubMed: 14766359]
2. Stein D, Wu J, Fuqua SA, Roonprapunt C, Yajnik V, D'Eustachio P, Moskow JJ, Buchberg AM, Osborne CK, Margolis B. The SH2 domain protein GRB-7 is co-amplified, overexpressed and in a tight complex with HER2 in breast cancer. *EMBO J.* 1994; 13(6):1331–1340. [PubMed: 7907978]
3. Tanaka S, Pero SC, Taguchi K, Shimada M, Mori M, Krag DN, Arai S. Specific peptide ligand for Grb7 signal transduction protein and pancreatic cancer metastasis. *J. Natl. Cancer Inst.* 2006; 98(7): 491–498. [PubMed: 16595785]
4. Siamakpour-Reihani S, Argiros HJ, Wilmeth LJ, Haas LL, Peterson TA, Johnson DL, Shuster CB, Lyons BA. The cell migration protein Grb7 associates with transcriptional regulator FHL2 in a Grb7 phosphorylation-dependent manner. *J. Mol. Recognit.* 2009; 22(1):9–17. [PubMed: 18853468]

5. Ivancic M, Daly RJ, Lyons BA. Solution structure of the human Grb7-SH2 domain/erbB2 peptide complex and structural basis for Grb7 binding to ErbB2. *J. Biomol. NMR.* 2003; 27(3):205–219. [PubMed: 12975581]
6. Spuches AM, Argiros HJ, Lee KH, Haas LL, Pero SC, Krag DN, Roller PP, Wilcox DE, Lyons BA. Calorimetric investigation of phosphorylated and non-phosphorylated peptide ligand binding to the human Grb7-SH2 domain. *J. Mol. Recognit.* 2007; 20(4):245–252. [PubMed: 17705331]
7. Porter CJ, Matthews JM, Mackay JP, Pursglove SE, Schmidberger JW, Leedman PJ, Pero SC, Krag DN, Wilce MC, Wilce JA. Grb7 SH2 domain structure and interactions with a cyclic peptide inhibitor of cancer cell migration and proliferation. *BMC Struct. Biol.* 2007; 7:58. [PubMed: 17894853]
8. Ambaye ND, Pero SC, Gunzburg MJ, Yap M, Clayton DJ, Del Borgo MP, Perlmutter P, Aguilar MI, Shukla GS, Peletskaya E, Cookson MM, Krag DN, Wilce MC, Wilce JA. Structural basis of binding by cyclic nonphosphorylated peptide antagonists of Grb7 implicated in breast cancer progression. *J. Mol. Biol.* 2011; 412(3):397–411. [PubMed: 21802427]
9. Brescia PJ, Ivancic M, Lyons BA. Assignment of backbone ¹H, ¹³C, and ¹⁵N resonances of human Grb7-SH2 domain in complex with a phosphorylated peptide ligand. *J. Biomol. NMR.* 2002; 23(1):77–78. [PubMed: 12061724]
10. Ivancic M, Spuches AM, Guth EC, Daugherty MA, Wilcox DE, Lyons BA. Backbone nuclear relaxation characteristics and calorimetric investigation of the human Grb7-SH2/erbB2 peptide complex. *Protein Sci.* 2005; 14(6):1556–1569. [PubMed: 15930003]
11. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* 1998; 54(Pt 5):905–921. [PubMed: 9757107]
12. Xia B, Tsui V, Case DA, Dyson HJ, Wright PE. Comparison of protein solution structures refined by molecular dynamics simulation in vacuum, with a generalized Born model, and with explicit water. *J. Biomol. NMR.* 2002; 22(4):317–331. [PubMed: 12018480]
13. Chen J, Im W, Brooks CL 3rd. Refinement of NMR structures using implicit solvent and advanced sampling techniques. *J. Am. Chem. Soc.* 2004; 126(49):16038–16047. [PubMed: 15584737]
14. Linge JP, Williams MA, Spronk CA, Bonvin AM, Nilges M. Refinement of protein structures in explicit solvent. *Proteins.* 2003; 50(3):496–506. [PubMed: 12557191]
15. Lange OF, van der Spoel D, de Groot BL. Scrutinizing molecular mechanics force fields on the submicrosecond timescale with NMR data. *Biophys. J.* 2010; 99(2):647–655. [PubMed: 20643085]
16. Still WC, Tempczyk A, Hawley RC, Hendrickson T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* 1990; 112(16):6127–6129.
17. Eck MJ, Shoelson SE, Harrison SC. Recognition of a high-affinity phosphotyrosyl peptide by the Src homology-2 domain of p56lck. *Nature.* 1993; 362(6415):87–91. [PubMed: 7680435]
18. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J. Mol. Graph.* 1996; 14(1):33–38. 27–38. [PubMed: 8744570]
19. Hoof RW, Vriend G, Sander C, Abola EE. Errors in protein structures. *Nature.* 1996; 381(6580):272. [PubMed: 8692262]
20. Vriend G. WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* 1990; 8(1):52–56. 29. [PubMed: 2268628]
21. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics.* 2006; 22(2):195–201. [PubMed: 16301204]
22. Schirra HJ, Anderson MA, Craik DJ. Structural refinement of insecticidal plant proteinase inhibitors from *Nicotiana glauca*. *Protein Pept. Lett.* 2008; 15(9):903–909. [PubMed: 18991765]
23. Porter CJ, Wilce MC, Mackay JP, Leedman P, Wilce JA. Grb7-SH2 domain dimerisation is affected by a single point mutation. *Eur. Biophys. J.* 2005; 34(5):454–460. [PubMed: 15841400]
24. Rahuel J, Gay B, Erdmann D, Strauss A, Garcia-Echeverria C, Furet P, Caravatti G, Fretz H, Schoepfer J, Grutter MG. Structural basis for specificity of Grb2-SH2 revealed by a novel ligand binding mode. *Nat. Struct. Biol.* 1996; 3(7):586–589. [PubMed: 8673601]

25. Songyang Z, Shoelson SE, McGlade J, Olivier P, Pawson T, Bustelo XR, Barbacid M, Sabe H, Hanafusa H, Yi T. Specific motifs recognized by the SH2 domains of Csk, 3BP2, fps/fes, GRB-2, HCP, SHC, Syk, and Vav. *Mol. Cell Biol.* 1994; 14(4):2777–2785. [PubMed: 7511210]
26. Pero SC, Oligino L, Daly RJ, Soden AL, Liu C, Roller PP, Li P, Krag DN. Identification of novel non-phosphorylated ligands, which bind selectively to the SH2 domain of Grb7. *J. Biol. Chem.* 2002; 277(14):11918–11926. [PubMed: 11809769]
27. DeLano, WL. The PyMOL Molecular Graphics System. San Carlos, CA, USA: DeLano Scientific; 2002.
28. Ulrich EL, Akutsu H, Doreleijers JF, Harano Y, Ioannidis YE, Lin J, Livny M, Mading S, Maziuk D, Miller Z, Nakatani E, Schulte CF, Tolmie DE, Kent Wenger R, Yao H, Markley JL. BioMagResBank. *Nucleic Acids Res.* 2008; 36(Database issue):D402–408. [PubMed: 17984079]
29. Case, DA.; Darden, TA.; TE Cheatham, I.; Simmerling, CL.; Wang, J.; Duke, RE.; Luo, R.; Merz, KM.; Pearlman, DA.; Crowley, M.; Walker, RC.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, KF.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, DH.; Schafmeister, C.; Ross, WS.; Kollman, PA. AMBER 9. San Francisco: University of California; 2006.
30. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins.* 2006; 65(3):712–725. [PubMed: 16981200]
31. Onufriev A, Bashford D, Case DA. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins.* 2004; 55(2):383–394. [PubMed: 15048829]
32. Onufriev A, Bashford D, Case DA. Modification of the Generalized Born Model Suitable for Macromolecules. *J. Phys. Chem. B.* 2000; 104(15)(15):3712–3720.
33. Homeyer N, Horn AH, Lanig H, Sticht H. AMBER force-field parameters for phosphorylated amino acids in different protonation states: phosphoserine, phosphothreonine, phosphotyrosine, and phosphohistidine. *J. Mol. Model.* 2006; 12(3):281–289. [PubMed: 16240095]
34. Gordon JC, Myers JB, Folta T, Shoja V, Heath LS, Onufriev A. H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* 2005; 33(Web Server issue):W368–W371. [PubMed: 15980491]

Figure 1.

Refined solution structure of the Grb7-SH2 domain in complex with the pY1139 peptide.

(A) Overlay of the refined NMR ensemble, excluding disordered end segments. The backbone trace is shown, with residues 17–116 of the SH2 domain in black and residues 1139–1142 of the peptide in red. (B) Overlay of entire residue range for the refined ensemble. A single model is shown in smooth-loop cartoon representation (residues 17–116 and 1139–1142 only) and is overlaid with the backbone trace of all ten models of the ensemble. Disordered end regions are distinguished using blue for the domain N-terminus and green for the domain C-terminus as well as the peptide termini. Regions in the SH2 domain with substantial conformational variability across models are highlighted in orange. Other residues are displayed in black for the domain and red for the peptide, as in panel A.

Figure 2. Comparison of the orientation of the F90 and F99 sidechains in the dimeric, ligand-free Grb7-SH2 domain (panel A) and in the monomeric Grb7-SH2/pY1139 complex (panel B). The two structures are shown in approximately the same orientation, having been aligned prior to capturing the images. Displayed are SH2 domain residues 17–116 and, where applicable, peptide residues 1139–1142. **(A)** Single monomer of the X-ray crystal structure for the dimeric ligand-free Grb7-SH2 domain (2QMS chain A, Porter *et al.* 7)], shown as a cartoon drawing with the sidechains of F90 (F502) and F99 (F511) visualized. **(B)** Cartoon drawing of a representative model of the new Grb7-SH2/pY1139 ensemble, with residues F90 and F99 displayed as sticks for the representative model. The approximate positions of the sidechains of F90 and F99 in the remaining NMR models (after backbone structural alignment) are shown as lines. The pY1139 peptide is imaged using stick representation.

Figure 3.

The pY+3 hydrophobic binding pocket is exposed in the ligand-free Grb7-SH2 domain crystal structure (panel **A**) but occluded in the pY1139-bound NMR ensemble (panel **B**). (**A**) Crystallographic chain A of the ligand-free Grb7-SH2 domain (2QMS) [7], shown with opaque surface representation in the upper image and with transparent surface and solid cartoon representations in the lower image. The solvent-exposed hydrophobic surface formed by the sidechains of F43, V45, Y68, I70, M83, V97, I106, and L107 is highlighted in purple (sidechains also displayed for these residues in the lower image). (**B**) A representative NMR model of the Grb7-SH2/pY1139 complex (new ensemble), showing the same hydrophobic residues in purple. The solvent-exposed surface of the hydrophobic patch formed by these residues is diminished. Residues D84 and D85 of the EF loop largely occlude the pY+3 binding pocket and are shown in yellow, including their sidechains in stick representation (lower image). The pY1139 peptide is displayed in pink, with its sidechains shown in line representation (lower image). Figures rendered with PyMOL [27].

Table 1

Quantitative Comparison of the Currently Reported (“New”) and Previously Reported [5] (“Original”) Ensembles for the Grb7-SH2/pY1139 Complex

NMR ensemble	New	Original
Violation statistics ^a		
Average number of		
Distance restraint violations > 0.5 Å	0.0	0.4
Distance restraint violations > 0.2 Å	4.8	7.7
Dihedral angle restraint violations > 5 °	1.5	<i>not reported</i>
Maximum distance restraint violation	0.40 Å	<i>not reported</i>
Maximum dihedral angle restraint violation	7.9°	<i>not reported</i>
RMSD to ensemble average coordinates ^b		
<i>SH2 domain residues 17–116 and pY1139 peptide residues 1139–1142</i>		
Backbone atoms	0.99 Å	0.71 Å
All atoms	1.57 Å	1.23 Å
RMSD from idealized covalent geometry ^c		
Covalent bonds	0.016 Å	0.008 Å
Angles	2.7°	1.1°
PROCHECK-NMR Ramachandran plot statistics ^c		
<i>For entire ensemble (10 models)</i>		
<i>Excluding glycine, proline, phosphotyrosine, and end residues</i>		
Residues in most favored regions	69.8 %	55.6 %
Residues in additional allowed regions	28.0 %	37.1 %
Residues in generously allowed regions	2.1 %	6.6 %
Residues in disallowed regions	0.1 %	0.6 %
WHAT IF statistics ^d		
<i>For a single NMR model, before and after its refinement</i>		
Z-scores, indicating the number of standard deviations		
from the expected value for well-refined X-ray structures:		
Second generation packing quality	−3.87	−4.74
Ramachandran plot appearance	−0.76	−0.96
<i>chi-1/chi-2</i> rotamer normality	−0.64	−0.72
Backbone conformation	−5.83	−6.68
RMS Z-scores, expected to fall around 1.0:		
Bond lengths	0.80	0.36
Bond angles	1.55	0.62
Inside/Outside residue distribution	1.16	1.18

^aViolation statistics for the original ensemble as reported previously [5].

^bValues for both ensembles determined using the RMSD Calculator Tool in VMD [18].

^cStatistics acquired using the ADIT validation server of the Protein Data Bank.

^dWHAT_CHECK (WHAT IF) assessment performed *via* the Swiss Model Workspace [19–21].