Multiscale Conformational Heterogeneity in Staphylococcal Protein A: Possible Determinant of Functional Plasticity

Lindsay N. Deis,1 Charles W. Pemble IV,2,3 Yang Qi,1 Andrew Hagarman,1,4 David C. Richardson,1 Jane S. Richardson,1 and Terrence G. Oas1,*

1Department of Biochemistry, Duke University, Durham, NC 27710, USA
2Human Vaccine Institute, Duke University School of Medicine, Durham, NC 27710, USA
3Macromolecular Crystallography Center, Duke University, Durham, NC 27710, USA
4Present address: KBI Biopharma, 1101 Hamlin Road, Durham, NC 27704, USA
*Correspondence: oas@duke.edu
http://dx.doi.org/10.1016/j.str.2014.08.014

SUMMARY

The Staphylococcus aureus virulence factor staphylococcal protein A (SpA) is a major contributor to bacterial evasion of the host immune system, through high-affinity binding to host proteins such as antibodies. SpA includes five small three-helix-bundle domains (E-D-A-B-C) separated by conserved flexible linkers. Prior attempts to crystallize individual domains in the absence of a binding partner have apparently been unsuccessful. There have also been no previous structures of tandem domains. Here we report the high-resolution crystal structures of a single C domain, and of two B domains connected by the conserved linker. Both structures exhibit extensive multiscale conformational heterogeneity, which required novel modeling protocols. Comparison of domain structures shows that helix1 orientation is especially heterogeneous, coordinated with changes in side chain conformational networks and contacting protein interfaces. This represents the kind of structural plasticity that could enable SpA to bind multiple partners.

INTRODUCTION

The structural plasticity conferred by conformational flexibility has increasingly been recognized as a likely determinant of function. For example, multiscale heterogeneity in the calmodulin central helix most likely helps it in binding >100 protein targets (Wilson and Brunger, 2000), and a concerted motion seen in both nuclear magnetic resonance (NMR) and crystal structures of ubiquitin is proposed to underlie its functional plasticity of promiscuous binding to many different proteins with high affinity (Lange et al., 2008). However, flexibility is manifested in a variety of ways, depending both on the protein itself and on how it is observed. Flexibility is apparent in X-ray crystallography as electron-density inconsistent with a single molecular model—either fully separated peaks or anisotropic density shapes showing fluctuation of atom groupings. In this work, we refer to alternative conformations as conformational heterogeneity rather than flexibility because the latter term implies motion on a relevant time scale, which cannot be determined by crystallography. Many phenomena contribute to conformational heterogeneity in crystal structures, from diverse crystal contacts to functionally relevant conformational fluctuations on a wide range of time and size scales.

Like ubiquitin, staphylococcal protein A (SpA) exhibits broad binding specificity with other proteins. This protein allows Staphylococcus aureus to evade the innate and adaptive immune systems, making it a significant challenge to human health. Among virulence factors responsible for S. aureus pathogenicity, SpA is the best studied and arguably the most important. It is a highly abundant 42kDa multi-domain cell-surface polypeptide with two functionally distinct halves (Figure 1A). The C-terminal half anchors SpA to the extracellular surface of the peptidoglycan cell wall via the LPXTG motif (Schneewind et al., 1992) and is likely disordered due to its low sequence complexity. In contrast, the N-terminal half is a series of five stable protein-binding domains (E-D-A-B-C). Recent studies establish that the conserved sequence KADNKF forms a highly flexible linker between all domains except E to D, which uses the longer sequence KADAQQNKF, also likely to be highly flexible (A.H. and T.G.O, unpublished data). The five domains have sequence identities of 74% to 91% (relative to A domain; Figure S1 available online) and share the same three-helix-bundle topology. The folding of each domain is thermodynamically uncoupled to the others and displays a gradient of increasing stability toward the more C-terminal modules (A.H. and T.G.O, unpublished data). In addition, the B domain rapidly unfolds and refolds approximately 70 times per second (Myers and Oas, 2001), and recent studies establish the same property for the other four domains (A.H. and T.G.O, unpublished data). All five domains can bind the Fc domain with Fc (Protein Data Bank [PDB] 1FC2) (Deisenhofer, 1981) and D domain with Fab (1DEE) (Graille et al., 2000). These
structures show partner interactions with SpA, but they lack comparison with unbound domain structures, and their lower resolution (2.7–2.8 Å) does not allow determination of multiple conformations. The B domain/Fc cocrystal structure (1FC2) lacks coordinates for most of helix3, which originally stimulated conformations. The B domain/Fc cocrystal structure (1FC2) interhelical angle between helix1 and helices2 and 3 (Gouda NMR solution structures of B domain and its variants is the observed in the Fc complex (Deisenhofer, 1981), whereas a larger angle was found in the original solution structures (Gouda et al., 1992; Tashiro et al., 1997). However, the differences in helix1-2 angle between most structures are not significantly larger than the uncertainties in these most recent structures, so the existence and possible role of a coordinated conformational change in helix1, whose residues form the majority of the contacts with Fc, remains unresolved.

Conformational heterogeneity is traditionally evaluated at two scales: side-chain-rotamer conformations and tertiary-structure rearrangements, such as hinge motions and secondary-structure reorientations. It includes local backbone differences such as peptide flips or backrubs that accompany side-chain rotamer sampling (Davis et al., 2006), small translational shifts of a few residues, and also side chain changes between distinct rotamer states, which involve large movements and significant energy barriers. Large-scale heterogeneity involves concerted positional differences of many residues, including changes in helix-helix positioning (Zheng et al., 2004) or β sheet twist and rare cases of major refolding (Skehel and Wiley, 2000), as well as the long-recognized interdomain hinge motions. Any of these types of conformational dynamics can be integral to enzyme function, allosteric regulation, induced-fit binding and functional plasticity, by enabling the alternative structures required for biological function.

To investigate the molecular basis for SpA flexibility and the connection between local and global conformational heterogeneity more generally, we have determined several X-ray crystal structures of SpA C domain and B-B (two B domains connected by the conserved linker), all in the absence of any partner protein. C domain was solved to 0.9 Å resolution at cryogenic temperature (Figure 1B). The structure shows many backbone and side chain alternative conformations. Because previous work suggested that cryogenic freezing limits protein conformational heterogeneity (Fraser et al., 2011; Juers and Matthews, 2004), we also determined a 1.4 Å structure of C domain at room temperature. The B-B construct, determined at 1.5 Å resolution (Figure 1C), is an informative mimic of the nearly identical A-B or B-C pairs. Because the linker was fully visible, the B-B structure reveals effects of the special case of linker proximity on conformational heterogeneity. This work establishes that SpA protein-binding domains exhibit extensive structural plasticity that presumably helps enable a small domain to accommodate multiple binding partners, and sheds light on the diverse nature of that plasticity.

RESULTS

Overview of C Domain and B-B Structures

Diffraction data for P2₁ C-domain crystals were collected at cryogenic and room temperatures and solved by single-wavelength anomalous dispersion (Table 1). Both structures contain one molecule in the crystallographic asymmetric unit and include all 58 residues of the SpA C domain (Figure 1B), plus a Zn²⁺ ion that binds the chain ends. Both also share the same three-helix-bundle topology seen in previous SpA-domain structures (Figure 1B). Including N- and C-caps (Richardson and Richardson,
Table 1. Data Collection and Refinement Statistics

<table>
<thead>
<tr>
<th></th>
<th>C domain (Cryogenic)</th>
<th>C domain (Room Temperature)</th>
<th>B-B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁</td>
<td>P2₁</td>
<td>P6₃</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td>27.3, 38.4, 28.6</td>
<td>27.8, 38.6, 29.0</td>
<td>44.4, 44.4, 214.8</td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>90.0, 117.4, 90.0</td>
<td>90.0, 118.2, 90.0</td>
<td>90.0, 90.0, 120.0</td>
</tr>
<tr>
<td>Wavelength</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50.0–0.9 (0.92–0.9)</td>
<td>50.0–1.42 (1.44–1.42)</td>
<td>50.0–1.49 (1.52–1.49)</td>
</tr>
<tr>
<td>Rsym</td>
<td>0.07 (0.20)</td>
<td>0.07 (0.33)</td>
<td>0.10 (0.62)</td>
</tr>
<tr>
<td>l/c</td>
<td>33.7 (4.36)</td>
<td>29.7 (2.54)</td>
<td>28.0 (2.2)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>97.5 (77.6)</td>
<td>97.0 (73.6)</td>
<td>99.9 (99.6)</td>
</tr>
<tr>
<td>Redundancy (%)</td>
<td>6.1 (2.7)</td>
<td>5.6 (2.9)</td>
<td>6.8 (4.4)</td>
</tr>
<tr>
<td>FOM</td>
<td>0.84</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>25.39–0.90</td>
<td>25.5–1.42</td>
<td>37.9–1.49</td>
</tr>
<tr>
<td>No. reflections</td>
<td>37,938</td>
<td>9,889</td>
<td>39,026</td>
</tr>
<tr>
<td>Rwork/Rfree</td>
<td>11.3/13.0</td>
<td>11.4/14.3</td>
<td>14.2/18.5</td>
</tr>
<tr>
<td>No. atoms</td>
<td>Protein 1,646</td>
<td>1,524</td>
<td>6,240</td>
</tr>
<tr>
<td>Zinc</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Acetate</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>134</td>
<td>50</td>
<td>409</td>
</tr>
<tr>
<td>B-factors</td>
<td>Protein 7.4</td>
<td>16.4</td>
<td>19.8</td>
</tr>
<tr>
<td>Zinc</td>
<td>17.8</td>
<td>12.7</td>
<td>–</td>
</tr>
<tr>
<td>Acetate</td>
<td>7.3</td>
<td>16.4</td>
<td>–</td>
</tr>
<tr>
<td>Water</td>
<td>19.6</td>
<td>36.3</td>
<td>32.5</td>
</tr>
<tr>
<td>Root-mean-square deviations</td>
<td>0.013</td>
<td>0.013</td>
<td>0.003</td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>1.554</td>
<td>1.412</td>
<td>0.712</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>Model validation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramachandran outliers (%)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ramachandran favored (%)</td>
<td>98.2</td>
<td>99.0</td>
<td>99.0</td>
</tr>
<tr>
<td>Rotamer outliers (%)</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>C-beta outliers (%)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Clashscore</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Overall score</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
</tr>
</tbody>
</table>

aData were collected from a single crystal. Values in parentheses are for the highest-resolution shell.

bFrom the MolProbity-style validation in Phenix.

Overview of SpA Conformational Heterogeneity

The high-resolution data show that SpA domains have much more conformational heterogeneity than observed in the previous crystal structures, or even than other crystal structures generally. Discrete backbone alternative conformations were identified for 55% of residues (31 of the 58 total) in room-temperature C domain, 62% in cryogenic C domain, 61% in B-B chain X, and 74% in B-B chain Y. Doing justice to such complexity required non-traditional approaches to achieve consistency within and between the alternative models, as described in the Experimental Procedures section. The extensive alternative-conformation modeling contributed to achieving the outstanding validation statistics reported in Table 1.

In addition to side chain shifts riding on the more global backbone shifts, all three structures have extensive individual side chain heterogeneity. Figure 2A shows examples for two surface-charged side chains and an interior aliphatic side chain, with electron density to support as many as four distinct rotamers. To quantify conformational heterogeneity among all six domains, we computed the maximum distance between equivalent atoms of alternative conformations within each residue along the sequence, separately for backbone atoms and for side chain atoms. The 3D distribution of this heterogeneity within single domains is shown for cryogenic C in Figure 2B as the local width in a putty-sausage diagram. The greatest side chain heterogeneity occurs at side chains pointing away from the interior, not unexpectedly. This surface heterogeneity occurs in side chains without strong intermolecular interactions, but an exception is the Tyr114 side chains in B-B chains X and Y, which interact with each other across molecules in the asymmetric unit.

Figures 2C, S3, and S4 use various representations to compare the distribution of conformational heterogeneity between the six different domains in these structures. The overall pattern is very similar in the C domains and the first domains of B-B. An interesting exception is the lower heterogeneity of C domain at the beginning of helix3, close to the main cluster of sequence differences between C and B. The most conspicuous contrast, however, is between domain2 of both B-B chains and that of the other four domains. In particular, helix1
is significantly less heterogeneous in domain2, probably due to interaction between the interdomain linker and helix1 of domain2 (discussed below). This perturbation by the linker appears to induce greater changes in conformational heterogeneity than data collection temperature, sequence differences between B and C domains, or the fact that C domain is 0.5 ± 0.15 kcal/mol more stable than B domain (A.H. and T.G.O, unpublished data).

Conformational Coordination of Side Chain Alternative Conformations

Many residue-level conformational changes involve rotamer shifts of neighboring side chains (Figures 3A–3C), where only some local combinations of conformations are possible without prohibitive steric clashes. Because electron density only shows the sum of all conformations, other information is needed to correctly analyze the conformational clustering, including all-atom sterics and the logic of assigning alternative conformation labels and occupancies.

The two domains within each B-B chain have nearly the same backbone conformation, but differ in backbone heterogeneity (see above) and also show differences in side chain conformations, whereas Trp113 is in turn limited by the presence of the linker bound to domain2.

Based on refined occupancies of Gln110 in chain Y and Tyr114 in both chains, we can estimate the populations of three sets of rotamers, which are depicted in Figures 3A–3C. The most populated set (Figure 3A) has a population of $\sim 60\%$, and the second set (Figure 3C) has a population of $\sim 30\%$. The least populated set (Figure 3B) has a much lower population of $\sim 8\%$, presumably because of a somewhat unfavorable contact between the Tyr rings, which may move apart by an amount too small to be distinguished within the summed density.

Heterogeneity in Helix1 Orientation

With the structures of unbound C domain and B-B solved, we now have six structures of SpA domains with which to analyze and compare conformational heterogeneity, as well as the various single-domain structures already in the PDB. Montelione and coworkers (Zheng et al., 2004) previously reported wide disagreement over the tilt angle of helix1. Our structures demonstrate that these differences represent actual differences in helix1 orientation between structures, rather than errors in one or more of the NMR structures. Our results confirm that helix2...
is the least variable and helix1 the most, with more change parallel than perpendicular to the helix2-3 plane. To allow quantitative comparison of helix1 orientations among current SpA-domain structures, we transformed the helix axis vectors for each domain into the same coordinate system with helix2 as the principal reference (see Experimental Procedures). The superimposed helix axes and the relative helix1 orientations for ten different SpA-domain structures are shown in Figure 4A and Table S1, respectively. Our crystal structures differ by up to 7.5° from one another (Figure 4B), 10° from 1Q2N (Zheng et al., 2004), and 13° from 1DEE (Graillé et al., 2000).

Although there are some sequence differences between the different domains, these nonconserved residues are not at the helical interfaces, so differences in interhelical angles cannot be attributed directly to local sequence differences. In addition, nearly the full range is seen among B-B domains. We hypothesize that subtle differences in packing due to residue-level conformational differences in and out of the direct interface are responsible for helix1 rearrangements in single-domain SpA structures. To test this, we compared the interhelical packing in C domain with the canonical B-domain NMR structure, 1Q2N (Zheng et al., 2004; Figure 4C). Helix1 differs most toward its C-terminal end, with a pivot point around Ile16. In analogy to Crick’s “knobs and holes” (Crick, 1953), this important “knob” side chain nests into a deep cavity between helix2 and helix3 and makes six side chain contacts across the interface (Figure 4D). Ile16 adopts the same conformation in both structures, and four of those six interface contacts (to Phe30, Ile31, Leu45, and Ala48) are identical. This is consistent with previous work (Braithed and Wells, 1996), which suggested that Ile16 stabilizes the packing of helix1 to helix2. The interaction of Leu45 with Ala12 is also the same in both structures. Figure 4D shows the key knob residues on the N- and C-terminal ends of helix1 that intercalate with “hole” residues across the interface, thereby holding the three-helix bundle together. Most assume unique conformations in each structure, leading to differences in knob locations (Figure 4D) when viewed down the helical axis. For example, both Leu19 and Asp52 sidechains shift significantly to maintain their contact, while Leu22 changes rotamer from mt (minus, trans) to tp (trans, plus) in most 1Q2N models to maintain contact with Ile19 in the new helix1 orientation.
consistently shifted toward the surface of the complex, leaving gaps in the packing.

Measurement of interfacial solvent-accessible surface area (Chothia, 1976) is a useful way to correlate structure to binding affinity for evolved, well-packed interfaces, but ignores gaps of up to 2.8 Å. As a more sensitive measure for this modeled and possibly ill-packed interface, we used all-atom contact analysis (Word et al., 1999), which counts only atom-atom contacts within 0.5 Å of ideal van der Waals distance. The cognate interface in 1FC2 has 142 Å² total all-atom contact, 110 Å² of favorable H-bond plus van der Waals, and one clash (see Table S1 for details). The initial noncognate interface has only 54 Å² favorable and 12 clashes; the rebuilt noncognate interface has only one clash but still just 68 Å² of favorable contact. For comparison, the internal interface between helix1 and 2-3 is 153 Å² total (147 Å² favorable), the B-B linker contacts domain2 across 106 Å², and a sample of tight biological dimer and inhibitor interfaces ranged from 100 to 460 Å². Crystal contacts typically have fairly little direct contact—a sample ranged from 18 to 40 Å². Of note for SpA, however, C domain makes crystal contact across 121 Å² plus a Zn site, and the B-B crystals alternative contacts of 144 and 258 Å²; presumably this helps attain high resolution.

The coordinated network of changes in side chain conformation found in the noncognate modeling is recapitulated elsewhere in our crystal structures, and it gives a broader context for the constraints on possible rotamer pairings seen in Figure 3 when assigning consistent alternative-conformation models in the B-B structure. Of our six domain structures, the high interhelix angle and side chain arrangement of the C domain is also seen in the second domains of B-B, whereas the low interhelix angle and side chain arrangement of 1FC2 B domain is seen in the first domains of B-B. Notably, all six domains make extensive crystal contacts across this helix1-2 surface, in three different arrangements. As shown in Figures 5C and 5D, seven side chains make concerted rotamer changes: Ile31 on helix2 and Gin10, Asn11, Phe/Trp13, Tyr14, Leu17, and His18 on helix1. Some of the coordination between adjacent side chain rotamers is constrained but not all: the more exposed Asn11 and His18 adopt multiple alternative conformations, and are not sequence conserved. In summary, there is substantial coordination between backbone and side chain conformational heterogeneity in SpA domains, and a probable relationship between that combined heterogeneity and the ability to form contacts with a variety of protein partners.

DISCUSSION

Interpreting Conformational Heterogeneity in Crystallographic Data

Representing protein structures as collections of possible models consistent with the data is common practice for structures determined by NMR. Although these are often called ensembles, there is usually no direct evidence that they represent the true ensemble of structures present in solution. Recent efforts to incorporate dynamic NMR information into NMR model refinement have produced more realistic representations of true ensembles (Lindorff-Larsen et al., 2005). Ensemble building is a relatively new pursuit in X-ray crystallography, and three current methodological approaches are promising but still in their infancies. The first approach repeats crystallographic model building and refinement from multiple randomly seeded starts (DePristo et al., 2004; Terwilliger et al., 2007), which is good at identifying uncertainty within the major conformation but seldom captures large differences or minor populations. The second approach is molecular dynamics-based ensemble refinement, for instance as currently under development in the PHENIX crystallographic suite (Adams et al., 2010; Burnley et al., 2012). When we tried it on our structures, individual results failed validation and only some of the side chain alternative conformations evident in electron density were captured. A third approach is
High Angle

B-factors until all visible backbone and side chain alternative density (see Experimental Procedures). Briefly, we kept isotropic extended traditional modeling of discrete alternative conformations one residue at a time. To accomplish this, we limited to careful identification of alternative conformations. Therefore, we used manual placement particularly that of the backbone, is difficult for the current automated methods to fit. However, the subtle anisotropy observed in our electron density, up the tedious process of modeling side chain heterogeneity with calculated rotamer changes in a network of seven side chains (bold colors) for (C) two low helix-angle structures (B-B domain 1 for both chains), similar to the B domain in 1FC2, and for (D) two high helix-angle structures (C domain and B-B chain Y Domain2). Switching only side chain network conformations on the superimposed, high-angle, noncognate C model from (B) can eliminate most clashes but neither recreates good van der Waals packing nor any H-bonds. However, the high-angle domain structures each form an extensive, well-packed contact with another SpA-domain partner in the crystal. See also Table S1.

Figure 5. Coordination of Backbone and Side Chain Conformational Changes with Binding of Fc

(A and B) All-atom contacts at binding interfaces. (A) Well-fit, cognate interface of Fc with SpA-B domain (low-angle helix 1-2 orientation), in PDB file 1FC2; (B) Poorly fit, noncognate interface of Fc with superimposed high-helix-angle conformation of C domain from 4NPD. Green and blue dots show good van der Waals contact or H bonds, and clusters of red spikes show steric clashes. (C and D) Correlation of changes in helix-helix angle with concerted rotamer changes in a network of seven side chains (bold colors) for (C) two low helix-angle structures (B-B domain 1 for both chains), similar to the B domain in 1FC2, and for (D) two high helix-angle structures (C domain and B-B chain Y Domain2). Switching only side chain network conformations on the superimposed, high-angle, noncognate C model from (B) can eliminate most clashes but neither recreates good van der Waals packing nor any H-bonds. However, the high-angle domain structures each form an extensive, well-packed contact with another SpA-domain partner in the crystal. See also Table S1.

Figure 22, 1467–1477, October 7, 2014 ©2014 Elsevier Ltd All rights reserved

Conformational Heterogeneity of Protein A

the qFit program (van den Bedem et al., 2009), which uses low-level electron density to identify potential alternative conformations for the next atom but has not yet been extended to model complete alternative side chain conformations or to correlate neighboring changes. Automated methods for modeling and coupling conformational networks, like that used in qFit and CONTACT (van den Bedem et al., 2013), respectively, speed up the tedious process of modeling side chain heterogeneity and very accurately couple side chain alternative conformations. However, the subtle anisotropy observed in our electron density, particularly that of the backbone, is difficult for the current automated methods to fit. Therefore, we used manual placement of sidechain and backbone atoms to identify all alternative conformations.

Because of these difficulties, our initial steps toward building a complete, consistent conformational ensemble from the crystallographic data were limited to careful identification of alternative conformations one residue at a time. To accomplish this, we extended traditional modeling of discrete alternative conformations based on anisotropy or separate peaks in the electron density (see Experimental Procedures). Briefly, we kept isotropic B-factors until all visible backbone and side chain alternative conformations had been built. These conformations were assigned to specific, internally consistent models of alternative conformations using a combination of MolProbity conformational and all-atom clash analysis, plus PHENIX refinement of occupancies.

Origin of High SpA-Domain Conformational Heterogeneity

Increasingly more structures in the PDB break 1 Å resolution, where alternative conformations are most visible. Even among these ultra-high-resolution structures, however, most have about 10% backbone heterogeneity; a few (e.g., 1EXR, 2116, 1YK4, 3NIR, and 2VB1) include alternative conformations that stretch for more than two residues and have between 25% and 36% backbone alternative conformations. We have found only the PDB ID 1M40 β-lactamase (Minasov et al., 2002), fit with 63% backbone alternative conformations, in or near the indeed unusual 55%–74% range seen in our SpA structures. The backbone alternative conformations within a single SpA-domain crystal structure differ by only 1 Å or less, but their pattern is very similar to the larger changes between structures. Prevalence of side chain alternative conformations in SpA is large but hard to compare quantitatively with other structures because many ride on backbone shifts.

There are several features of SpA that may lead to this unusual conformational heterogeneity. For example, SpA domains are small, single-chain three-helix bundles, each held together by the Ile16 pivot residue (Figure 4D). The lack of an extensive hydrophobic core might facilitate conformational heterogeneity by allowing adjustment of interhelical angles with less constraint from interior van der Waals interactions.

Another consideration is that multiple backbone conformations may be less rare in other proteins than suggested by their occurrence in crystal structures, because all types of partial disorder are more prevalent at low resolution but can be well characterized only at high resolution. An unusual feature of SpA domains may be that, despite having high conformational heterogeneity, they can still form high-resolution crystals. All our crystals have extensive contacts that form a continuous zigzag of domains along a crystallographic screw axis—a heterologous contact in P21 for C domain, and an alternation of domain1 and domain2 pseudo-2-folds in P65 for B-B, each of which uses the helix1-2 face. The overall domain orientations must repeat in exactly 180° for P21 and in exactly 60° for P65, which these domains achieve using varied interhelical angles and side chain positions. Therefore, we speculate that these well-ordered crystals are enabled by the same plasticity used to bind different biological targets.

Another unusual property of SpA domains is that they unfold and refold very rapidly, in less than 100 ms (Myers and Oas, 2001; A.H. and T.G.O, unpublished data). Although kinetic and thermodynamic experiments establish that the folding reaction is two-state, the rapid kinetics implies a relatively low activation barrier between the folded and unfolded ensembles. This characteristic of the overall energy landscape may be correlated with the complex native-state landscape suggested by the high conformational heterogeneity we observe. In any case, it seems likely that evolution has converged for functional reasons on a sequence that not only includes a flexible linker but also provides side chain networks that can support alternative helix packing (Figure 4) and alternative binding interfaces (Figure 5).
Another intriguing side chain-level feature of the current structures is a possible role for Phe5 in binding interactions that involve the helix1-2 surface. The Phe5 side chain flips out to allow contacts with a second molecule in the 1FC2, 1DEE, and 1H0T complexes. Given the rather low order parameter of its NH in solution (A.H. and T.G.O., unpublished data), Phe5 must spend considerable time disordered, along with the rest of the linker. However, it is tucked down over one end of the helix1-2 surface in all six of our SpA-domain structures and in the NMR structures 1SS1, 1BD0, 2SPZ, and 1Q2N (where NOEs were observed between the Phe ring and domain surface residues [Doreleijers et al., 2005]). The helix1-2 surface, the most hydrophobic face on SpA domains, forms three quite different, well-packed crystal interfaces in our structures, and is likely to play an important role in binding other biological partners of SpA besides Fc.

Most interestingly, the large total set of SpA domain structures now available, especially the ones at high resolution, allow study of possible coupling between intradomain conformational heterogeneity and the variety of binding partners. We have shown that: (1) helix1 orientation differs by 7.5° among our six domain structures, and up to 13° from other crystal and NMR structures; (2) helix orientation is correlated with a network of side chain rotamer changes on the helix1-2 face (Figure 5); and (3) Fc binding to that face is compatible only with the low-helix-angle form of the concerted side chain/helix rearrangement, whereas other well-packed contacts incorporate the high-angle version (Table S1). It is extremely difficult to establish that backbone conformational change is essential for binding many partners, but we can definitively say that concerted backbone/side chain changes are in fact used for that binding.

Multidomain Conformational Preferences in SpA

The contact between D domain and F_{ab} uses a different face of SpA domains than the contact between B domain and F_{c}. As a result, a single domain could actually bind both F_{ab} and F_{c} at the same time, which Graille et al. (2000) demonstrated by building a hypothetical model of the ternary complex. Whereas some side chain conformations and helix orientations found in C domain and B-B are compatible with binding F_{c} (as discussed above), the domain-domain relationship observed in each B-B chain would block binding of F_{c} to either domain due to major steric overlap with the other linked domain (Figure 6C). This observation confirms the NMR evidence that, in solution, B-B is free to adopt many other domain-domain conformations, thus allowing F_{c} to bind. Unlike the case with F_{c}, the global conformation of the B-B chain is compatible with binding up to two F_{ab} molecules simultaneously on the negatively charged helix2-3 faces, with no steric interference (Figure 6D). Cocryostals of B-B or other SpA multidomain constructs with F_{c} or F_{ab} would be desirable to understand SpA binding.

ExPERIMENTAL PROCEDURES

Plasmid Construction

The C-domain gene was PCR cloned from the SpA-N gene. The PCR primers added a 5’ Ndel and a 3’ BamHI site, and were subsequently cloned into the T7 expression plasmid pAE4D (Doering, 1992). The B-B gene was synthesized by GENEWIZ in a pUC57 cloning vector, and was subsequently cloned into the pAE4D expression vector. It contained the F13W substitution to aid detection of the protein.
**Structure**

**Conformational Heterogeneity of Protein A**

**Protein Expression and Purification**
Plasmids were transformed into *Escherichia coli* BL21(DE3) cells using standard transformation procedure. A single colony of transformed bacteria was used to inoculate a 50 ml culture of Luria broth media with 0.1 mg/ml ampicillin. This starter culture was incubated at 37°C until the optical density (OD) reached 0.8–1.0, whereupon it was used to inoculate 1 l cultures that were allowed to grow to OD 0.8–1.2 at 37°C. Isopropyl-beta-D-thiogalactopyranoside was then added to a final concentration of 1 mM, and the cultures were incubated for an additional 6–8 h. The cells were harvested by centrifugation and resuspended in 20–30 ml of 50 mM Tris pH 8.8, 1 mM EDTA, and protease inhibitor cocktail (AEBSF, pepstatin, bestatin and E-64). The cells were lysed in a French pressure cell, and insoluble material was centrifuged from the lysate. The lysate was brought to pH 9.0, and micrococcal nuclease was added to digest large DNA fragments for 15 min. The resulting solution was brought to 4 M guanidinium HCl (BioBasic) and 20 mM TCEP was added. In two successive steps, the solution was dialyzed in a 5% acetic acid buffer, which precipitated many cellular materials, but not expressed proteins. After centrifugation of the insoluble material, the resulting solution was allowed to dialyze overnight into deionized water. The protein was further purified using two types of ion-exchange chromatography. First, the protein was loaded onto a strong cation exchanging, SP Sepharose (GE) column in 50 mM acetate buffer at pH 3.6. The column was eluted with a NaCl gradient (typically 100 mM to 500 mM NaCl) in a volume of 600–800 ml and collected in 10 ml fractions monitored by a UV detector (Bio-Rad) at 287 nm. The fractions comprising the protein elution peak were checked for purity by SDS-PAGE. The purest fractions were pooled and dialyzed against deionized water. Subsequently, the resulting solution was loaded onto a weak anion exchanging DEAE Sephacel (GE) column in 25 mM Tris (BioBasic) at pH 10.0 and eluted with an 800 ml NaCl gradient (typically 0–250 mM gradient) into 10 ml fractions monitored by a UV detector at 287 nm. The protein elution peak fractions were checked for purity by SDS-PAGE. The purest fractions were pooled and dialyzed against deionized water. The final solution was lyophilized and stored in a desiccator. Purities of the final protein stocks were confirmed by SDS-PAGE to be >95% pure. The masses of all proteins were confirmed with electrospray ionization mass spectrometry.

**Crystallization, Data Collection, and Initial Models**
C-domain protein solution was prepared at 20 mg/ml from lyophilized protein and mixed in a 1:1 ratio with crystallization solution (55 mM ZnCl₂, 1.6% PEG 6000, and 0.1 M MES pH 6). Crystals formed by hanging-drop vapor diffusion within five days and were of the space group P4₁2₁2₁, although not of high-quality. Seed stock was prepared by adding approximately 10 of the P4₁2₁2₁ crystals into a 50 µl aliquot of well solution in a Hampton Seed Bead, and the resulting microseeds were used in a new crystal screen. Diffraction-quality crystals were formed by hanging-drop vapor diffusion within 2 days by mixing C-domain well solution (30 mM NaSCN, 26% [w/v] PEG 3,350, and 0.5% [w/v] glycerol), and seeds in a ratio of 3:2:1, respectively.

For the cocrystallization data set, cocrystallization of the crystal was achieved with the addition of 30% glycerol to the crystallization condition, and the crystal was frozen by direct immersion in liquid nitrogen. All data were collected remotely at the Advanced Photon Source (APS) at Argonne National Laboratory, beamline 22-ID (SER-CAT). Data were processed and scaled with HKL2000 (Otwinowski and Minor, 1997; Table 1).

The B-B structure was solved by molecular replacement in PHASER (McCoy et al., 2007) using the cocrystallized C-domain structure as the search model, which has a 91% sequence similarity to each domain of B-B. The search model was modified to remove any alternative conformations and ten mobile residues at the N terminus. The initial model for the individual domains of B-B was built by AUTOBUILD (Terwilliger et al., 2008) and then rebuilt in COOT (Emsley et al., 2010). The residues in the interdomain linkers were manually built in COOT (Emsley et al., 2010).

**Refinement and Heterogeneous Model Building**
Refinement of all models was carried out in phenix.refine (Adams et al., 2010). Initial refinement was performed using two isotropic B factors per residue. Once the model was placed, it was refined using isotropic Bs for all atoms (Figure S2A). After several rounds of iterative refinement and model building, the visible concerted anisotropy was used to fit main chain alternative conformations in segments (e.g., helix1). Each alternative-conformation segment was rotated and translated in COOT (Emsley et al., 2010) to fit the direction of anisotropy in the carbonyl oxygen 2Fo–Fc density, effectively separating these alternative conformations (Figure S2B). The side chain for each alternative conformation was then fit down to 0.3–0.5Å levels. Candidates that could not be fit rotationally or that later refined to zero occupancy were discarded. To avoid bad geometry across peptide bonds, backbone alternative conformations were continued to the flanking Cx atoms if needed, using backbrubs (Davis et al., 2009) in a new PHENIX (Adams et al., 2010) utility written for this purpose.

After all main chain and side chain alternative conformations were fit, several rounds of occupancy and anisotropic B-factor refinement were performed. All atoms but H were refined as anisotropic for the cocrystallized C-domain structure, and all atoms but H and waters were refined as anisotropic for the room-temperature and B-B structures. The 2Fo–Fc and Fo–Fc density at each residue was checked after each round of refinement. Final refinement for all structures was carried out in phenix.refine using automatic weight optimization. The final refined models for the cocrystallized structure and room-temperature structure had Rwork/Rfree values of 11.0%/12.7% and 11.2%/14.1%, respectively, whereas the final refined model for the B-B structure had Rwork/Rfree values of 14.2%/18.5%.

**Illustrations**
Coordinate files for molecules in crystal contact were produced in Chimera (Pettersen et al., 2004). Their all-atom contact dots were produced by the interface feature in MolProbity (Chen et al., 2010), counted in Mage (Richardson and Richardson, 2001), and converted to square anstroms by dividing by the default contact-dot density of 16Å². Figure 5 was made in KINO (Chen et al., 2009). Figure 4A was constructed using Prekin and Mage. All remaining structural illustrations were made in PyMOL (http://www.pymol.org/). Figure 2C was made in Mathematica (2010).

**ACCESSION NUMBERS**
The Protein Data Bank accession numbers for the molecular coordinates and structure factors reported in this work are 4NPD (cocrystallized C-domain), 4NPE (room temperature C domain), and 4NPB (B-B).

**SUPPLEMENTAL INFORMATION**
Supplemental Information includes Supplemental Experimental Procedures, four figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.str.2014.08.014.

**AUTHOR CONTRIBUTIONS**


ACKNOWLEDGMENTS

This paper is dedicated to the memory of the late Prof. Thomas Alber, whose vision led us to see proteins in a new and illuminating way. Crystal screening and data processing were conducted in collaboration with the Duke Macromolecular X-ray Crystallography Shared Resource. Diffraction data were collected remotely at the Southeast Regional Collaborative Access Team 22-ID beamlines at the Advanced Photon Source, Argonne National Laboratory, supported by the US Department of Energy Sciences under contract number W-31-109-Eng-36. The plasmid construction, protein expression, and purification aspects of this work were supported by NIH grant R01-GM081666 (to T.G.O.). The crystallographic and structure analysis aspects of this work were supported by NIH grant R01-GM73930 and GM073919 (to D.C.R.) and in part by NIH grant P01-GM063210 Project IV (to J.S.R.). We thank Nigel Moriarty for the utility in PHENIX that spreads backbone alternative conformations to the flanking Cα atoms.

Received: June 24, 2014
Revised: August 12, 2014
Accepted: August 17, 2014
Published: October 7, 2014

REFERENCES


