BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

*Biochemistry* **2010,** *49,* 8689–8699 8689
DOI: 10.1021/bi100398p

# Three-Dimensional Structures Reveal Multiple ADP/ATP Binding Modes for a Synthetic Class of Artificial Proteins[†,‡]

C. R. Simmons,[§] C. L. Magee,[∥] D. A. Smith,[§,∥] L. Lauman,[∥] J. C. Chaput,[§,∥] and J. P. Allen*[,∥]

[§]*Center for Evolutionary Medicine and Informatics, The Biodesign Institute, Arizona State University, Tempe, Arizona 85287, and* [∥]*Department of Chemistry and Biochemistry, Arizona State University, Tempe, Arizona 85287*

*Received March 15, 2010; Revised Manuscript Received July 29, 2010*

ABSTRACT: The creation of synthetic enzymes with predefined functions represents a major challenge in future synthetic biology applications. Here, we describe six structures of de novo proteins that have been determined using protein crystallography to address how simple enzymes perform catalysis. Three structures are of a protein, DX, selected for its stability and ability to tightly bind ATP. Despite the addition of ATP to the crystallization conditions, the presence of a bound but distorted ATP was found only under excess ATP conditions, with ADP being present under equimolar conditions or when crystallized for a prolonged period of time. A bound ADP cofactor was evident when Asp was substituted for Val at residue 65, but ATP in a linear configuration is present when Phe was substituted for Tyr at residue 43. These new structures complement previously determined structures of DX and the protein with the Phe 43 to Tyr substitution [Simmons, C. R., et al. (2009) *ACS Chem. Biol. 4*, 649−658] and together demonstrate the multiple ADP/ATP binding modes from which a model emerges in which the DX protein binds ATP in a configuration that represents a transitional state for the catalysis of ATP to ADP through a slow, metal-free reaction capable of multiple turnovers. This unusual observation suggests that design-free methods can be used to generate novel protein scaffolds that are tailor-made for catalysis.

A major challenge of protein design and evolution is the creation of protein scaffolds that fold into stable three-dimensional structures with predefined functions. Of particular interest is the development of methodologies that harness the information content of biological macromolecules to facilitate the emergence of synthetic machines that function as efficient catalysts in future biotechnology, environmental, and medical applications (*1*, *2*). Computational methods provide one approach to this challenge by designing amino acid sequences that are compatible with a target structure or that introduce active site residues into an existing protein scaffold (*3*). Notable experimental accomplishments in this area include the design of an entirely new protein that folds in a novel topology (*4*), as well as the redesign of natural proteins into active catalysts (*5*−*7*). Despite such progress, the design of tailor-made enzymes remains difficult as the catalytic activities of synthetic proteins remain much more limited than those of natural enzymes. Achieving improved functionality will likely require new strategies that reduce or eliminate many of the evolutionary constraints associated with natural proteins. One possibility is to apply the principles of Darwinian evolution to large pools of sequences as a means of identifying individual proteins that fold into arbitrary structures with predefined functions (*8*). This approach, commonly known as de novo protein evolution, eliminates many of the biases associated with proteins that have evolved or been designed from natural scaffolds, leading to a more complete understanding of the folds that are possible in the protein universe.

The de novo approach raised the question of how frequently a functional protein might occur in a pool of random sequences, and this led to the discovery that proteins with ligand binding activity are sufficiently common in protein sequence space that they can be discovered using purely stochastic methods. Keefe and Szostak created the first de novo evolved proteins by isolating four synthetic ATP binding proteins from a pool of 4 trillion sequences of unique 80-mers (*9*). The three-dimensional structure of a high-affinity variant, identified as 18-19, was determined by X-ray crystallography (*10*, *11*) and revealed a novel $\alpha/\beta$-fold with a unique sequence and topology (Figure 1). In addition to the bound nucleoside, the protein has a Zn cofactor that is thought to have a role in establishing and maintaining the fold. Because this protein arose entirely through an in vitro process that randomly sampled different regions of sequence space, the structure suggested that the protein universe might contain many additional examples of proteins whose folds are physically realistic but for historical reasons, or perhaps for reasons of fitness, lie outside of the boundary that nature has defined as biological space.

Using directed evolution, we have since optimized protein 18-19 for enhanced stability (*11*−*13*). One of these stability-enhanced proteins, identified as DX, has the same overall fold as 18-19 despite amino acid differences at residues 32 and 65 relative to 18-19 (Figure 1). A surprising feature of the three-dimensional structures of the 18-19 and DX proteins was the presence of well-defined electron density for most of the bound adenosine cofactor,

[†]This work is supported by a grant from the National Science Foundation (MCB0821032).

[‡]The atomic coordinates and structure factors of the 1:1 DX, 1:10 DX-3 day, 1:10 DX-21 day, 1:1 Y43F, 1:1 D65V, and 1:100 D65V structures have been deposited in the Protein Data Bank as entries 3LTD, 3LTC, 3LTB, 3LTA, 3LT9, and 3LT8, respectively.

*To whom correspondence should be addressed: Department of Chemistry and Biochemistry, Arizona State University, Tempe, AZ 85257. Phone: (480) 962-8241. Fax: (480) 965-2747. E-mail: jallen@asu.edu.


Published on Web 09/07/2010

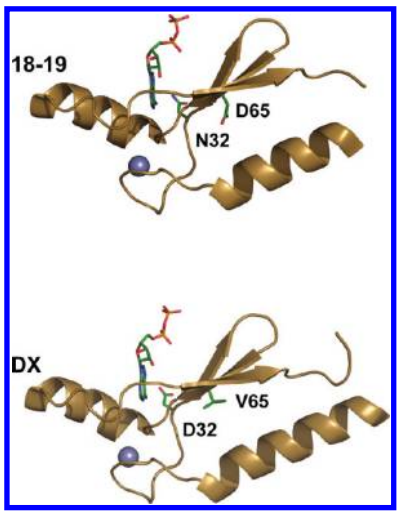


FIGURE 1: Ligand binding site of 18-19 and DX proteins. Ribbon model representation of the 18-19 and DX proteins with a bound ADP shown as a stick model as previously reported (*11*). The residues that differ between the two proteins, residues 32 and 65, are displayed with side chains. The ADP ligand and residue side chains are represented as stick drawings with atom-type colors (green for carbon, red for oxygen, blue for nitrogen, and orange for phosphate). The nonactive Zn metal ion is shown as a purple sphere. Single-letter amino acid notation was used as follows: D for Asp, V for Val, and N for Asn (Protein Data Bank entries 2P05 and 2P09 for 18-19 and DX, respectively).

including the α- and β-phosphates, but very weak density for the γ-phosphate, implying that the terminal phosphate was highly disordered or not present for unknown reasons (*11*). To improve our understanding of this phenomenon, we crystallized protein DX, under saturating amounts of ATP (*14*). The protein crystals obtained under these conditions diffracted to high resolution and yielded a new nucleotide-bound structure in which ATP was present and adopted a bent conformation that was held in place by a key water molecule that formed a bridge between basic residues in the binding site and the γ-phosphate of ATP. A null mutant identified as Y43F that contained a critical change of Tyr 43 to Phe in the binding site confirmed the role of the bridging water molecule in maintaining the bent geometry (*14*).

In this paper, a series of crystallographic experiments are reported with the aim of investigating the correlation between the ligand conformation and the functional properties of the protein. One outstanding unresolved question remaining was whether the protein DX was performing single- or multiple-turnover catalysis. To address this point, the structure of the DX protein was determined after exposure to ATP for different time periods and different ATP concentrations. Previously, the DX structure was determined, on the basis of the structure of the 18-19 protein, at limiting or saturating concentrations of added ATP and saturating concentrations of ADP (*11*, *14*). To minimize any model bias, the DX structures were determined using phases determined independently based upon multiple anomalous dispersion from the zinc cofactor. Because the two previous structures had been determined at only two extreme ATP concentrations, crystals were grown under intermediate ATP conditions and again under the limiting ATP condition. Also, a third DX structure was determined to address the aspect that in previous structural work exposure of the protein to ATP was performed for only a limited time. To examine the possibility of a time dependence of the ligand conformation and loss of the terminal phosphate, the DX structure was determined after exposure to ATP for an extended period of time.

In addition to the DX protein, structures are reported for two proteins containing single-site alterations relative to DX, Y43F and D65V. The Y43 position was chosen because the side chain was critical for the stability of a water molecule bound near the adenosine cofactor and the D65 residue appeared to play a critical role in the proposed enzymatic abilities of DX. Previously, the structure of Y43F had been determined only at a saturating concentration of ATP. To complement the studies on the DX protein, the structure of the Y43F protein was determined at limiting ATP concentrations. Biochemical studies had shown that the enhanced stability and other functional differences between DX and 18-19 are primarily due to the presence of Val 65 instead of Asp, with the second alteration of Asp 32 to Asn making a smaller contribution (*11*). How the structural differences at these two positions give rise to the observed functional differences was difficult to interpret. Therefore, the structure of D65V was determined, under both limiting and excess ATP conditions. Together, these six new structures provide a structural basis for interpreting the observed functional differences. The structural differences of the bound cofactor found in these structures are discussed in terms of the biophysical properties and the inferred enzymatic activity of these de novo proteins. Also discussed are the implications concerning the properties of ATP binding proteins found in biological systems.

## MATERIALS AND METHODS

*Protein Crystallization*. Proteins were expressed, purified, and crystallized using the sitting drop vapor diffusion method at room temperature as described previously (*11*−*14*). The protein was concentrated to ~20 mg/mL in 0.1 M sodium phosphate (pH 8.5), 0.4 M sodium citrate, 0.3 M sodium chloride, and 1 mM ATP, which represents an equimolar concentration. A protein drop with a volume of 1.5 μL was combined with an equivalent volume of reservoir solution containing 0.1 M sodium phosphate (pH 8.5), 0.4 M sodium citrate, 0.1−0.25 M sodium chloride, and 19−24% polyethylene glycol 400. Crystals grown under the saturation conditions at protein:ATP ratios of 1:10 and 1:100 were obtained by including 10 and 100 mM ATP, respectively, in the crystallization solution. We mounted all crystals directly from the crystallization solution within 3−4 days after setting up the crystallization trays and cryocooled them by plunging each crystal mounted in a loop into a liquid nitrogen bath. The exception was for the DX crystals grown under the 1:10 conditions for 21 days that were left in the crystallization solution for 21 days, after which crystals were looped and directly cooled in a 100 K nitrogen gas stream at the Arizona State University X-ray Diffraction Facility.

*Crystallographic Data Collection*. Data for the 1:10 crystals grown for 21 days were collected at the Arizona State University X-ray Diffraction Facility using an RAXIS IV++ detector at the Cu Kα absorption edge of 1.54 Å. All other data were collected at the NSLS X12b beamline using an ADSC Q4 CCD detector. Multiwavelength anomalous dispersion data for the 1:1 and 1:10 DX crystals were collected at the Zn absorption peak, inflection, and high-energy remote wavelengths ($\lambda$ = 1.2817, 1.2822, and 1.275 Å, respectively) based on a fluorescence scan. The 1:1 Y43F and D65V and 1:100 D65V data were collected at the Zn peak ($\lambda$ = 1.2817 Å). All crystals belonged to space group $P3_22_1$ and were isomorphic with the DX and 18-19 crystals (*11*−*14*). The diffraction data were indexed, refined, integrated, and scaled using the HKL2000 package (*15*). During data

processing, the unit cell parameters were refined to allow for any slight changes in crystal packing.

*Structure Determination and Refinement*. For the 1:1 Y43F, 1:10 Y43F, 1:1 D65V, and 1:100 D65V crystals, as well as the 1:10 DX crystals grown for 21 days, the structures were determined by difference Fourier analysis with the original DX structure (*11*) [Protein Data Bank (PDB) entry 2P09] stripped of all solvent, ligands, and cofactors, as the initial model. The 1:1 DX structure and the 1:10 DX structure grown for 3 days were determined using multiwavelength anomalous dispersion phasing from the zinc cofactor. The two Zn sites contained in the asymmetric unit (one site per protein molecule) were determined using the Python-based Hierarchical Environment for Integrated Xtallography (PHENIX) (*16*). The initial Zn substructure was determined using AutoSol; the AutoBuild program SOLVE was used for refinement and phase calculation, and the initial model was automatically built using RESOLVE (*17*). Manual model building for all structures was performed in Coot (*18*) with iterative rounds of refinement being performed using REFMAC5 from the CCP4 suite of programs (*19*) until the model was complete. The data were subsequently refined with the peak wavelength data. Calculations of $R_{free}$ for all six data sets were based upon 5% of the unique reflections. During further refinement, the resulting difference maps were utilized to determine the presence and configuration of either ATP or ADP at the active site, and the ligand was built into the electron density accordingly. Solvent molecules were built into regions having an appropriate chemical environment containing $F_o - F_c$ and $2F_o - F_c$ densities using cutoffs of $3.5r_{rms}$ and $1.0r_{rms}$, respectively. Additional rounds of model building and refinement were completed using REFMAC5 (*20*). The completed models were subsequently analyzed with Sfcheck and Procheck (*21*, *22*). All figures were generated with PYMOL (*23*). The Ramachandran plots of the models had 84.1−96.9% of the residues in the most favorable region of the plot, 3.1−12.7% in the additionally allowed region, and Lys 27 in the generously allowed region. In several models, Asn 37 lies in a disallowed region, but this residue is nevertheless contained within continuous electron density. This residue lies on a loop connecting the first and second β-strands that is disordered, and the poor geometry probably is reflective of the unresolved disorder. The models do not contain any cis peptides in the polypeptide chain.

## RESULTS

To probe the factors that influence the properties of the bound adenosine cofactor, structural studies of the three mutants, DX, Y43F, and D65V, were performed under different crystallization conditions. For each protein, hexagonal crystals, belonging to space group $P3_22_1$, appeared within 1 day, growing to a maximal size of ~0.1 mm. Each mutant was crystallized in the presence of equimolar (1 mM) and excess (10 or 100 mM) ATP and each structure determined independently. Normally, crystals were frozen after growing for 3 days, although in one set of experiments the DX crystals were allowed to remain in the crystallization solution for 21 days before being frozen for structural studies. The six structures determined from crystals grown under these different conditions are described below.

*Structures of the DX Protein*. Three crystal structures of DX were determined in the presence of 1 and 10 molar equivalents of ATP. All crystals were isomorphic and contained two DX monomers in the asymmetric unit. To minimize any model bias, the structures with 1 mM (1:1) and 10 mM (1:10) ATP were determined using multiple anomalous dispersion from the zinc cofactor. The third structure was determined from crystals grown in the presence of 10 mM ATP (1:10) that were undisturbed for 21 days before being frozen rather than the 3 day period used for all other crystals. This structure was determined using molecular replacement.

The 1:1 DX structure determined from crystals grown for 3 days in 1 mM ATP yielded a final refined model with $R_{cryst}$ and $R_{free}$ values of 18.0 and 21.8%, respectively, at 2.7 Å resolution; the 1:10 DX structure determined from crystals grown in 10 mM ATP for 3 days yielded a final refined model with $R_{cryst}$ and $R_{free}$ values of 17.9 and 21.1%, respectively, at 2.0 Å resolution, and the 1:10 DX structure determined from crystals grown in 10 mM ATP for 21 days was determined at a resolution limit of 2.6 Å ($R_{cryst}$ = 18.3%, and $R_{free}$ = 24.1%) from data collected on our in-house X-ray source (Table 1). For all three structures, a total of 69 of the 81 residues in the protein (residues 5−73) were well-defined in the $2F_o - F_c$ electron density maps at the $1.0\sigma$ level and are included in the final models. The 1:1 and 1:10 DX structures grown for 3 days are in very good agreement with one another with an overall root-mean-square deviation (rmsd) of 0.3 Å for the α-carbons. The DX structures have the same general fold as the previously reported structure of 18-19, with an rmsd of 0.5 Å for the α-carbons, except for some prominent differences occurring at residues Asn 32 and Asp 65 that are different amino acid residues in the two proteins (*11*).

The electron density determined from DX crystals grown in 1 mM ATP for 3 days yielded a well-ordered and unambiguous structure of ADP, with the adenine nucleobase well-coordinated via hydrophobic base stacking interactions of the aromatic side chain of Phe 50 and the phenolic ring of Tyr 43, as well as with main chain interactions with Gly 63 and Met 45 (Figure 2a). The α- and β-phosphates of the ADP molecule are coordinated in a concerted fashion with the polar side chains of Lys 34, Arg 41, and Tyr 43 and a series of solvent molecules. In contrast, the electron density determined for the DX structure in 10 mM ATP was consistent with the adenosine cofactor being ATP (Figure 2b). The ATP conformation is unusual as it is bent with the same coordination found for the ADP in the 1:1 DX structure at the nucleobase, but the ribose sugar and α-, β-, and γ-phosphates deviate significantly in their coordination scheme from typical ATP binding conformations observed in nature (*24*). One notable difference between DX crystallized with a 10-fold molar excess of ATP, compared to that with equimolar ATP, is the presence of a clear alternate conformation with an occupancy of 0.5 for the Asp 59 carboxylate side chain that has a position approximately 2 Å closer to the ATP binding pocket. In addition, the terminal nitrogen of the side chain of Lys 34 moves ~2 Å from the binding pocket when the protein is crystallized in the presence of 10 mM ATP. This structure also contained clear density for an alternate conformation for Asn 37 with an occupancy of 0.5. No notable differences of any other side chains were observed that could not be simply attributable to surface disorder.

The crystals grown for 21 days in 10 mM ATP yielded positive difference density that was described entirely by ADP (Figure 2c). Some residual positive $F_o - F_c$ density remained within the active site that could not definitively account for solvent or alternative nucleotide binding modes but was indicative of some ligand binding heterogeneity; however, the dominant configuration of ADP yielded an occupancy of 0.75 for the α- and β-phosphates. In this structure, the side chain nitrogen of Lys 34 is shifted ~2 Å into the binding pocket to coordinate the bound ADP. The refined model

Table 1: Crystallographic Data and Refinement of the DX Structures

| ligand (crystal drop) | 1:1 ATP | 1:10 ATP | 1:10 ATP |
|---|---|---|---|
| time of growth (days) | 3 | 3 | 21 |
| ligand (observed) | ADP | ATP | ADP |
| | Data Collection | | |
| wavelength (Å) | 1.2817, 1.2822, 1.275[e] | 1.2817, 1.2822, 1.275[e] | 1.5418 |
| space group | $P3_221$ | $P3_221$ | $P3_221$ |
| *a* (Å) | 73.15 | 73.70 | 72.26 |
| *b* (Å) | 73.15 | 73.70 | 72.26 |
| *c* (Å) | 54.80 | 54.76 | 54.88 |
| oscillation range (deg) | 0.5 | 0.5 | 0.5 |
| resolution range (Å) | 50−2.8 | 50−2.0 | 20−2.6 |
| total no. of observations | 24543, 22966, 21508[e] | 67477, 67480, 67554[e] | 30021 |
| no. of unique reflections | 7493, 7419, 6701[e] | 21937, 21985, 21945[e] | 5288 |
| multiplicity | 3.3, 3.1, 3.2[e] | 3.1, 3.1, 3.1[e] | 5.6 |
| $I/\sigma$ | 7.1, 7.0, 6.6[e] | 10.8, 11.0, 9.6[e] | 9.7 |
| $R_{sym}$[a] (%) | 14.8, 16.2, 17.0[e] | 7.3, 7.2, 8.9[e] | 13.4 |
| $R_{sym}$[a] (%) (last shell) | 50.5, 51.2, 40.7[e] | 37.5, 36.4, 43.6[e] | 55.3 |
| completeness (%) | 94.2, 91.6, 91.7[e] | 99.9, 99.9, 99.9[e] | 99.8 |
| | Refinement | | |
| $R_{cryst}$[b] (%), $R_{free}$[c] (%) | 18.0, 21.8 | 17.9, 21.1 | 18.3, 24.1 |
| rmsd[d] for bonds (Å) | 0.014 | 0.016 | 0.014 |
| rmsd[d] for angles (deg) | 1.742 | 1.672 | 1.699 |
| residues modeled | 5−73 | 5−73 | 5−73 |
| average *B* factor (Å$^2$) | 26.57 | 20.47 | 25.47 |

[a] $R_{sym} = \sum_h\sum_i[|I_i(h) - \langle I(h)\rangle|)/\sum_h\sum_i I_i(h)$, where $I_i(h)$ is the *i*th intensity measurement and $\langle I(h)\rangle$ is the weighted mean of all measurements of $I(h)$. [b] $R_{cryst} = \sum_h||F_o| - |F_c||/\sum_h|F_o|$, where $R_{cryst}$ is evaluated by the summation of all reflections used in refinement and $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes, respectively. [c] $R_{free} = \sum_h||F_o| - |F_c||/\sum_h|F_o|$, where $R_{free}$ is evaluated by randomly choosing 5% of the diffraction data not included in refinement. [d] Root-mean-square deviation from ideal values. [e] Peak, inflection, and remote, respectively.

was in good agreement with the 1:1 DX-3 day model with a calculated Cα rmsd of 0.2 Å.

*Structures of the D65V Protein*. The structure of D65V was determined in the presence of 1 mM ATP (1:1) at a resolution limit of 2.7 Å, with $R_{cryst}$ and $R_{free}$ values of 18.7 and 23.9%, respectively (Table 2). The resulting $2F_o - F_c$ and $F_o - F_c$ electron density maps showed the presence of one molecule of ADP that was built into a region of continuous $2F_o - F_c$ density (Figure 2d). For crystals grown in 100 mM ATP (1:100), the structure was determined at a resolution limit of 2.55 Å, with $R_{cryst}$ and $R_{free}$ values of 17.8 and 23.9%, respectively. In this case, the electron density yielded a refined ADP molecule in a configuration similar to that of the 1:1 condition (Figure 2e). The 1:1 and 1:100 D65V models are essentially identical to the 1:1 DX-3 day model with Cα rmsd values of 0.3 and 0.2 Å, respectively. The configuration of Asn 32, which is an Asp in the DX protein, is identical for both D65V structures, and no other notable differences involving side chains were observed.

*Structure of the Y43F Protein*. The structure of Y43F grown in 1 mM ATP (1:1) was determined at a resolution limit 2.7 Å, with $R_{cryst}$ and $R_{free}$ values of 20.2 and 23.8%, respectively (Table 2). The resulting $2F_o - F_c$ and $F_o - F_c$ electron density maps showed the presence of one molecule of ATP that was built into a region of continuous $2F_o - F_c$ density (Figure 2f). The ligand adopts an extended conformation that is much more conventional than that found in the 1:10 DX-3 day structure. Aside from the Y43F mutation, the overall Cα model is very comparable to the 1:1 DX structure with an rmsd of 0.3 Å. Compared with DX, the carboxylate of Asp 32 in the Y43F model is shifted ~1.5 Å from the nucleobase, and there does appear to be some heterogeneity at the Lys 34 position. The structure of the Y43F protein in excess ATP was not determined as this structure had been previously reported (*14*).

## DISCUSSION

In this paper, we report the determination of six structures of three related proteins, DX, Y43F, and D65V, under different crystallization conditions for the investigation of the ATP binding properties of a nonbiological protein that had been optimized for protein stability and ligand binding. These experiments were designed to study the possibility that bound ATP in the DX protein would undergo a slow catalytic conversion to ADP with subsequent release and repeated turnovers. To address this question, the DX protein was crystallized under three conditions: 1 mM ATP for 3 days, 10 mM ATP for 3 days, and 10 mM ATP for 21 days. The resulting structures are very similar with the same overall fold; however, the nature of the bound adenosine ligand differs significantly for these three conditions (Figure 3). When exposed to 1 mM ATP, the bound ligand is observed to be ADP. Increasing the ATP concentration from 1 to 10 mM for the normal 3 day period resulted in the presence of ATP, but increasing the crystallization time from 3 to 21 days results in the presence of ADP. The difference between these outcomes suggests that DX is a very slow catalyst that requires an unusually long incubation time because this protein was never optimized for catalysis and the structure does not contain a divalent metal ion in the active site. Because the half-life for ATP in our crystallization liquor is no less than ~12 months (*14*, *25*), we expected that ADP would be observed in the 21 day structure only if the protein was capable of multiple-turnover catalysis. Indeed, the structure of the 1:10 DX-21 day protein was found to contain ADP rather than ATP (Figure 3). In addition to these results for the DX protein, experiments were performed on two proteins that have single-site changes relative to DX, namely, D65V, which has Asp at position 32 rather than Asn, and Y43F, which has a Phe in place of Tyr at position 43. Under both equimolar

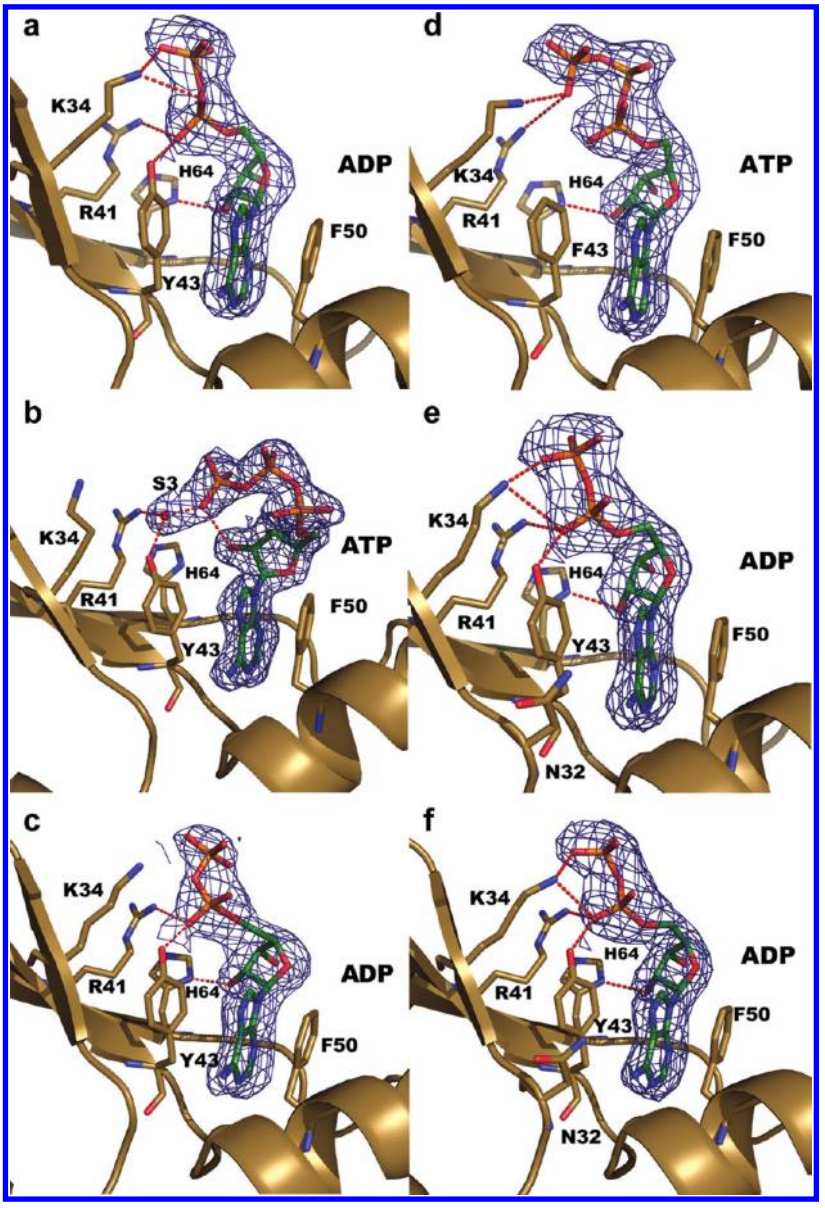


FIGURE 2: Electron density and three-dimensional structures of the DX, Y43F, and D65V proteins showing the bound cofactors and surrounding protein. Shown are (a) equimolar ATP growth conditions for the DX protein showing a bound ADP, (b) a 10-fold ATP excess with DX protein for 3 days showing a bound ATP in a highly bent conformation, (c) a 10-fold ATP excess with DX protein incubated for 21 days with a bound ADP, (d) equimolar ATP Y43F protein shown to have bound ATP, (e) equimolar ATP D65V protein with a bound ADP, and (f) a 100-fold ATP excess with D65V with a bound ADP. The $2F_o - F_c$ electron density maps are shown at $1.0\sigma$ except for that in panel b, which is contoured at $1.5\sigma$. Proteins are represented as ribbon diagrams with only the amino acid side chains that participate in ligand binding being displayed as stick models. Side chain atoms are colored by atom type (sand for carbon, red for oxygen, and blue for nitrogen). Bound ADP/ATP ligands are colored by atom type (green for carbon, red for oxygen, blue for nitrogen, and orange for phosphate).

and excess ATP conditions, the D65V protein has ADP present while Y43F has ATP present. Thus, the nature of the bound adenosine ligand is found to be dependent upon several factors, including the residues forming the ligand binding site, the concentration of ATP, and the time of exposure to ATP. Below we discuss in detail the relationship between these factors and the nature of the bound ligand, and the implications for cofactor binding in biological proteins.

*Comparison of Cofactor Structures*. The structures of four different proteins, 18-19, DX, D65V, and Y43F, have been determined under different crystallographic conditions using both multiple anomalous dispersion and molecular replacement. In all cases, the backbones of these structures are essentially identical, except for small differences associated with differences in the amino acid composition. In contrast to the distinct conservation of the protein backbones, many different configurations of the cofactors and binding site are observed. In four cases, the bound cofactor is ADP despite the addition of ATP to the crystallization solution, namely, for the DX protein crystallized under 1:1 and 1:10 21 day conditions and the D65V protein crystallized under 1:1 and 1:100 conditions. A comparison of the cofactors from these four structures shows essentially identical configurations of ADP and the surrounding protein (Figure 4). This configuration of ADP also matches that previously found in the structure of the DX protein grown using ADP (*14*). Thus, when ADP is present, the cofactor is well-defined and in a standard configuration that is the same under all crystallization conditions, for all three proteins, whether the ADP is directly added or ATP is added and then catalyzed into ADP.

In contrast to the situation for ADP, the two conditions under which ATP is found as the bound ligand, for 1:1 Y43F and 1:10 DX-3 day, yield two different configurations (Figure 4). In both cases, the adenosine components are nearly identical but an extended structure as commonly found in proteins is evident for 1:1 Y43F, while a highly distorted configuration is seen for the 1:10 DX-3 day structure. This pattern is also observed in the previously reported structures, with the differences between the 1:10 Y43F and 1:100 DX structures (*14*). The strong overlap of the adenosine moieties in all four structures is in agreement with the similar binding affinities of these proteins for ATP (*11*, *14*). The presence of ATP in the Y43F structure demonstrates a role of Tyr 43 in establishing conditions needed for catalysis, probably involving a water molecule that is hydrogen bonded to both Tyr 43 and the ATP cofactor.

Table 2: Crystallographic Data and Refinement of the D65V and Y43F Structures

| | D65V | D65V | Y43F |
|---|---|---|---|
| ligand (crystal drop) | 1:1 ATP | 1:100 ATP | 1:1 ATP |
| ligand (observed) | ADP | ADP | ATP |
| | Data Collection | | |
| wavelength (Å) | 1.2817 | 1.2817 | 1.2817 |
| space group | $P3_221$ | $P3_221$ | $P3_221$ |
| *a* (Å) | 70.70 | 71.46 | 71.11 |
| *b* (Å) | 70.70 | 71.46 | 71.11 |
| *c* (Å) | 55.94 | 55.58 | 55.94 |
| oscillation range (deg) | 0.5 | 0.5 | 0.5 |
| resolution range (Å) | 50−2.7 | 50−2.55 | 50−2.7 |
| total no. of observations | 23055 | 31012 | 19639 |
| no. of unique reflections | 4553 | 5562 | 4702 |
| multiplicity | 5.1 | 5.6 | 4.2 |
| $I/\sigma$ | 11.2 | 11.9 | 8.5 |
| $R_{sym}$[a] (%) | 10.1 | 9.2 | 13.5 |
| $R_{sym}$[a] (%) (last shell) | 51.6 | 49.4 | 57.0 |
| completeness (%) | 97.6 | 99.9 | 99.3 |
| | Refinement | | |
| $R_{cryst}$[b] (%), $R_{free}$[c] (%) | 18.7, 23.9 | 17.8, 23.9 | 20.2, 23.8 |
| rmsd[d] for bonds (Å) | 0.014 | 0.014 | 0.012 |
| rmsd[d] for angles (deg) | 1.746 | 1.71 | 1.69 |
| residues modeled | 5−73 | 5−73 | 5−73 |
| average *B* factor (Å$^2$) | 32.26 | 33.46 | 32.69 |

[a] $R_{sym} = \sum_h\sum_i(|I_i(h) - \langle I(h)\rangle|)/\sum_h\sum_i I_i(h)$, where $I_i(h)$ is the *i*th intensity measurement and $\langle I(h)\rangle$ is the weighted mean of all measurements of $I(h)$. [b] $R_{cryst} = \sum_h||F_o| - |F_c||/\sum_h|F_o|$, where $R_{cryst}$ is evaluated by the summation of all reflections used in refinement and $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes, respectively. [c] $R_{free} = \sum_h||F_o| - |F_c||/\sum_h|F_o|$, where $R_{free}$ is evaluated by randomly choosing 5% of the diffraction data not included in refinement. [d] Root-mean-square deviation from ideal values.

*ATP/ADP Ligand Binding Geometries*. To quantify the differences in the cofactor binding and explore the possibility that a de novo derived ATP binding protein might bind its ligand in a nontraditional manner as compared with known ATP binding protein structures (*24*), we examined the rotatable dihedral bonds from our six structures (Figure 5). Via comparison of these bound ligands, torsion angles 1—3 show relatively little deviation of the molecules from the favorable torsion angle regions that have been observed across a sampling of biological ATP binding proteins (*24*). For example, torsion angles 1−3 all have values within a narrow range. Despite the differences in the ATP conformations observed in the structures, ranging from an extended configuration for 1:1 Y43F to the highly distorted shape found in the 1:10 DX-3 day structure, there is little variation in the dihedral angles of the ligand when compared with that of a known protein that binds ATP in a bent conformation, namely aspartyl tRNA synthetase (*26*), or that of a protein that binds ATP in a more common extended conformation such as gluconate kinase (*27*). Across the glycosidic torsional angle, all modeled ligands adopt the most frequently observed *anti* conformer, which is clearly observable for torsion angle 1 (O4′−C1′−N9−C4), and are clustered closely in contrast to the variations of the angle for the ribose moiety (Figure 5).

In contrast, significant differences are observed upon comparison of torsion angles 4−6 for the ligands bound to de novo proteins and to biological proteins. The most notable deviations are found for the 1:1 Y43F ATP conformer, 1:10 DX bent ATP, and the bent ATP conformation observed in aspartyl tRNA synthetase. For example, the ATP conformation adopted by Y43F has values of 14.8° and —125.5° for torsion angles 4 and 5 (O3A−PA−O5′−C5′ and PB−O3A−PA−O5′), respectively, with these values being slightly outside the regions preferred by ATP in most

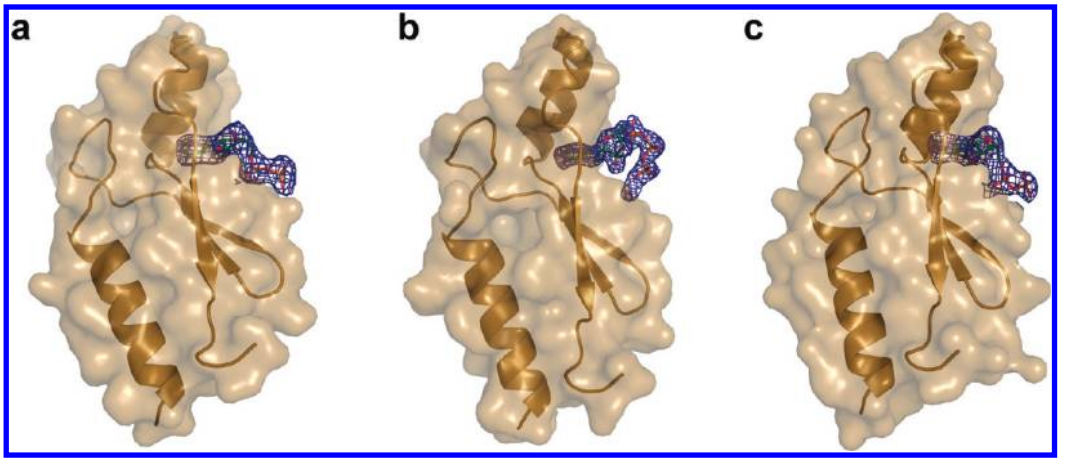


FIGURE 3: Comparison of the DX structures grown under different conditions. Three-dimensional structures of DX grown under (a) equimolar ATP conditions for 3 days showing a bound ADP, (b) 10-fold excess ATP conditions for 3 days with a bound ATP, and (c) 10-fold excess ATP conditions for 21 days with a bound ADP. Translucent surface representations are shown with a ribbon backbone (sand) with ATP/ADP ligands shown as stick models colored by atom type (green for carbon, red for oxygen, blue for nitrogen, and orange for phosphate). The $2F_o - F_c$ electron density maps (blue) are contoured at 1.0$\sigma$ in panels a and c and 1.5$\sigma$ in panel b.

proteins. Even with the drastically different conformations displayed by the ATP ligand for both the DX protein and aspartyl tRNA synthetase as compared with more common conformers, as exhibited for example by gluconate kinase, these ATP configurations share some significant deviations from the dihedral angles common to extended ATP conformers to facilitate the characteristic "bend" displayed by the molecule. As an example, torsion angle 4 has values of 102.57° and 52.6° in the 1:10 DX-3 day and aspartyl tRNA synthetase structures that significantly differ from the corresponding angle of −64.51° found for the extended configuration in gluconate kinase of extended ATP. Similarly, the distorted ATP configuration exhibited by 1:10 DX-3 day has a value of 102.5° for torsion angle 5 and adopts a turn that drastically differs from the dihedrals of all other bound ligands, which all have comparable angles. Although the ATP ligand in tRNA synthetase does display an unusual bent geometry, these proteins have never been reported to form any intramolecular polar interactions as observed in the 1:10 DX-3 day structure. Unlike the bent conformer in aspartyl tRNA synthetase, the γ-phosphate of the ATP in the 1:10 DX-3 day structure forms a strong 2.4 Å hydrogen bonding interaction with the 2′-OH of the ribose moiety. This is probably due to the fact that the most drastic torsion angle lies at position 5 on the phosphate chain instead of in position 4 in aspartyl tRNA synthetase, uniquely positioning it in a geometry that enables the polar contact.

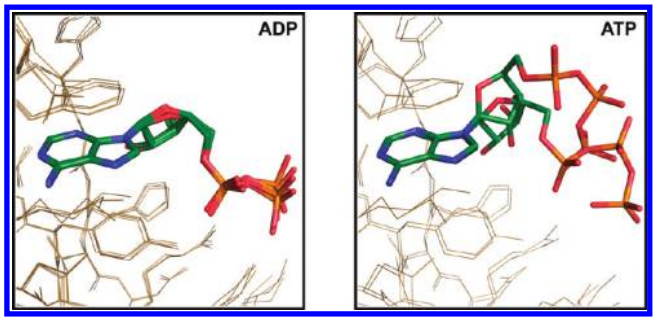


FIGURE 4: Comparison of ADP and ATP configurations. (a) Superposition of the ADP ligands observed under four conditions: equimolar ATP conditions for the DX protein, equimolar ATP conditions for the D65V protein, a 10-fold ATP excess for the DX protein grown for 21 days, and a 100-fold ATP excess for the D65V protein, colored by atom type (green for carbon, blue for nitrogen, red for oxygen, and orange for phosphate). (b) Superposition of the ATP ligands observed in the 10-fold ATP excess DX 3 day and equimolar Y43F structures (colored by atom type). Other atoms in the binding site are shown as line representations (brown).

*Role of Protein and Solvent Interactions in Cofactor Binding*. For both the ADP and ATP structures reported in this work, the adenosine moiety has strong stacking interactions with Tyr 43 and Phe 50 (Figure 6). Such hydrophobic base stacking

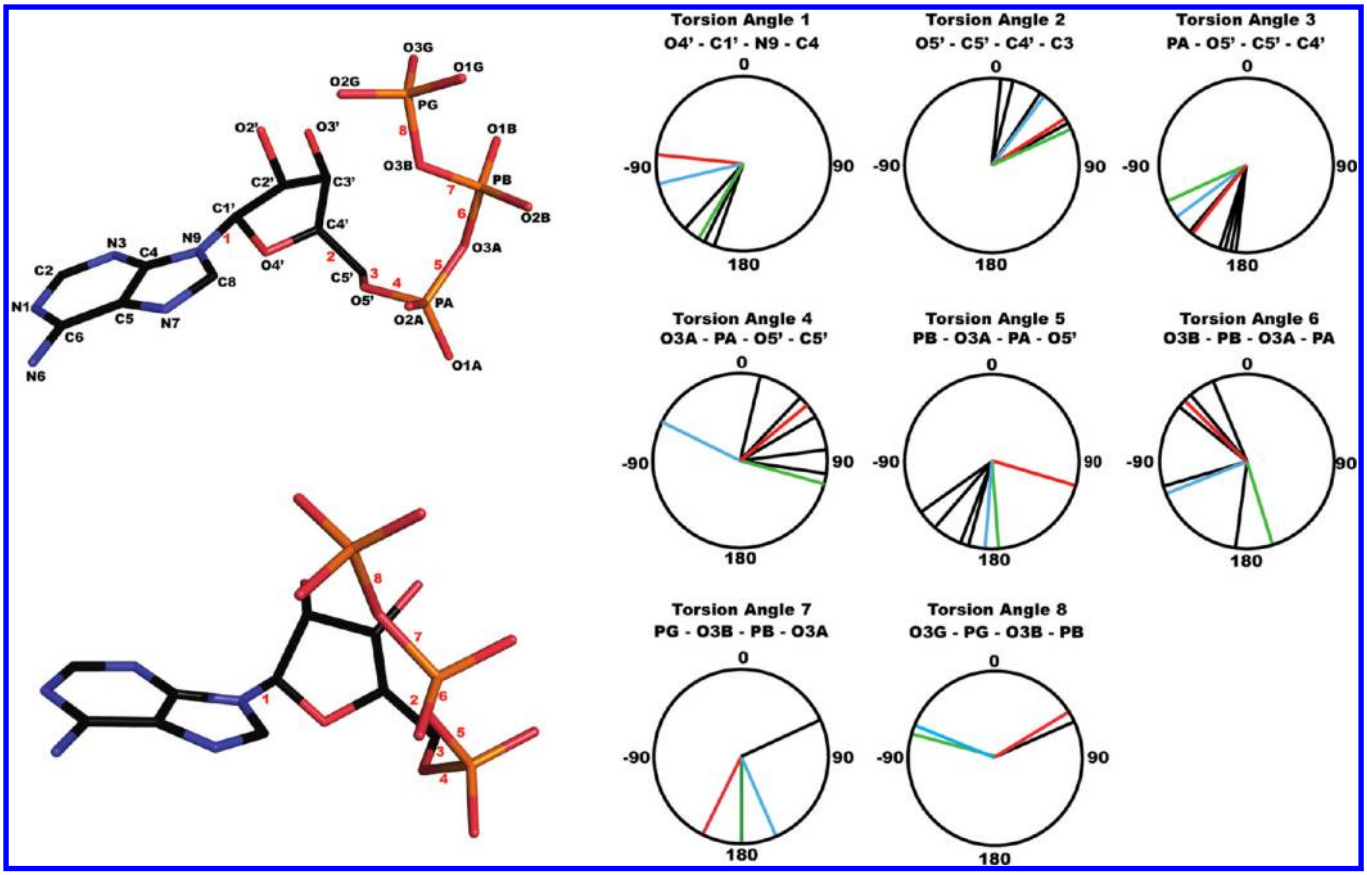


FIGURE 5: Comparative analysis of ATP/ADP ligand torsion angles. Stick model drawings show the comparison of ATP from a 10-fold ATP excess with DX and aspartyl tRNA synthetase, and gluconate kinase colored by atom type (black for carbon, blue for nitrogen, red for oxygen, and orange for phosphorus). Ligand atoms were labeled (in black) with the naming scheme from the deposited PDB files for each ATP ligand. Torsion angles were measured using Pymol for each rotatable bond in the ATP/ADP ligands, and the representative torsion angles are numbered in red for the eight rotatable bonds (ATP with eight torsion angles and ADP with six torsion angles). Angular wheel plots were generated for the measured torsion angles using SigmaPlot. The four atoms used to generate each torsion angle measurement are listed in the header for each wheel plot. Wheel plots 1−8 demonstrate (units of degrees) the comparison of torsion angles for DX bent (red radial lines) with ATP ligand conformations from aspartyl tRNA synthetase (*26*) (green radial lines) and gluconate kinase (*27*) (blue radial lines). The ligand torsion angles for the other four ADP molecules and the other ATP molecule presented in this work are plotted in black to provide a statistical frame of reference.

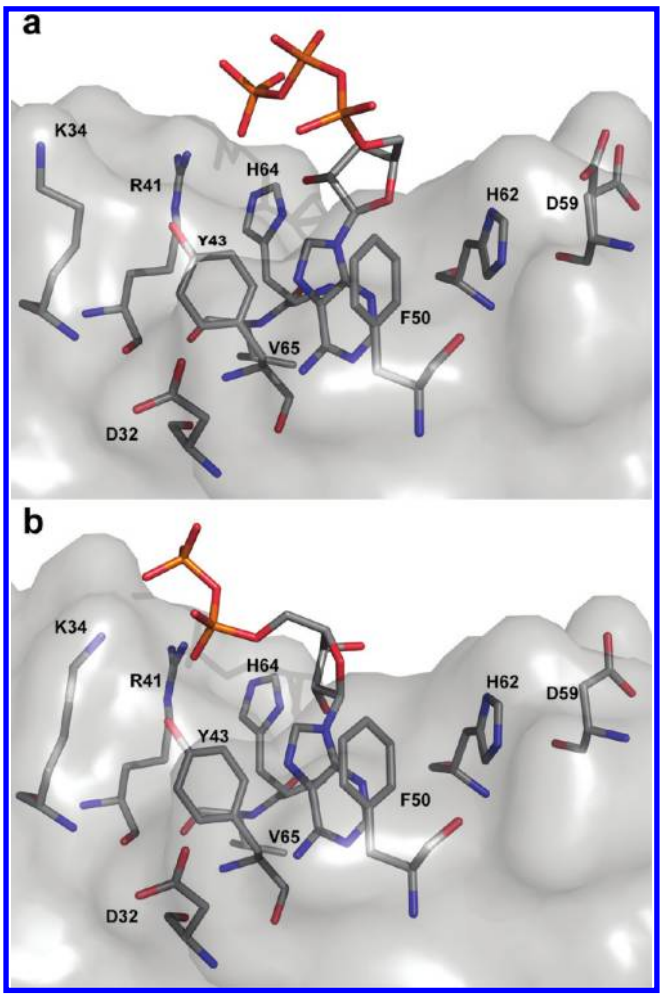


FIGURE 6: Structural landscape of the ATP and ADP binding domains. Shown are the cofactors, their binding regions, and translucent surfaces of DX grown in a 10-fold excess of ATP for (a) 3 days and (b) 21 days. Atoms are colored by atom type (gray for carbon, red for oxygen, blue for nitrogen, and orange for phosphate).

interactions with aromatic residues Phe, Tyr, and Trp are common in biological ATP binding proteins (*28*−*30*). For the bound ATP molecule in the 1:10 DX-3 day structure, the ribose sugar and the phosphate portions do not form any polar side chain contacts that have hydrogen bonding interactions with bound water molecules and an intramolecular hydrogen bond. In contrast, the structures with bound ADP demonstrate strong hydrogen bonding to the nearby polar residues Lys 34, Arg 41, Tyr 43, and His 64. Each of these residues is a member of the consensus sequence from the original selection (*11*) and may play a pivotal role in both binding and catalysis.

Nearly the entire sugar and phosphate chain of the ATP and ADP cofactors in these nonbiological proteins is solvent-exposed, unlike the case found in the majority of biological ATP binding proteins. In the 1:10 DX-3 day structure containing ATP bound in a bent conformation, the side chain of Asp 59 is present in a dual conformation (Figure 6). One conformation places the side chain carboxylate toward the binding pocket where, along with His 62, the residue serves to coordinate a solvent chain that tethers the bent ATP conformer to the protein. The alternate conformation with Asp directed away from the binding site is similar to what is observed when ADP is bound, so this conformation probably reflects a heterogeneity in which ADP is bound with low occupancy. Two residues, Asp 32 and Val 65, in tandem distinguish DX from 18-19 and influence the pronounced enhancement of folding stability and ligand binding of DX compared to that of 18-19 (*11*). Asp 32 resides close to the ATP binding site and is involved in coordinating a solvent molecule that interacts with the nucleobase. Val 65 is located on the periphery of the binding pocket.

Two of the high-resolution structures determined by our laboratory are the 1:10 DX-3 day structure reported here at 2.0 Å and the previously determined (*11*) 1:1 DX structure at a resolution limit of 1.65 Å (Figure 7). Both structures contain a well-ordered network of solvent molecules that are involved in ligand coordination, although these networks differ for the two cases. For the 1:10 DX-3 day structure, the distorted ATP is coordinated exclusively by a network of eight well-conserved solvent molecules and no side chain interactions. For the 1:1 DX structure, the bound ADP has five bound waters and five side chain contacts. The only two waters that are conserved between the two models (Figure 7) are identified as solvent S1, which coordinates the nucleobase, and solvent S2, which is involved in coordination of the α-phosphate of either ADP or ATP. The $\gamma$-phosphate of the ATP ligand in the 1:10 DX-3 day structure approaches the nucleoside binding pocket and forms a strong hydrogen bond (2.4 Å) interaction with the 2′-OH of the ribose sugar (Figure 7). In addition to this contact, the ligand is surrounded by a solvation sphere that forms hydrogen bonds that help to maintain the well-ordered but highly strained conformation. Water molecules on the backside of the bound ATP serve to coordinate the 2- and 3′-OH groups of the ribose, and each of the phosphate groups is contacted by a series of waters. Although none of the main or side chain atoms form contacts with the phosphate groups because of their protrusion into the solvent, the side chains of Lys 34, Arg 41, and Tyr 43 all serve to coordinate a single water molecule that forms a close hydrogen bond contact with the $\gamma$-phosphate of the ligand.

*Comparison to Biological Proteins*. All of the de novo evolved ATP binding protein structures with bound ADP molecules superimpose in virtually identical positions, while two different ATP configurations have been described (Figure 4). This linear configuration of ADP is the overwhelmingly preferred ligand conformation for the biological protein (*24*). In contrast, the distorted ATP configuration is rarely found in biological systems. In some cases, proteins have been observed to have ATP ligands that adopt more unusual conformations, as found in the class II tRNA synthetase family of proteins, like aspartyl and histidyl tRNA synthetase (*26*−*32*) in which the $\gamma$-phosphate is repositioned in near van der Waals proximity to the adenine moiety of itself, through an induced rotation of the phosphate chain. This class of proteins contains a conserved antiparallel $\beta$-sheet motif and requires this folding conformation, along with a concerted series of protein contacts, solvent molecules, and divalent metal cofactors to keep the ligand in position. In the case of the structure of ribonucleotide reductase, a similar bent conformation was observed, but it may be due to the fact that the crystals were soaked with ATP rather than the ligand being added directly to the mother liquor (*32*). Protein DX crystallized in the presence of a 10-fold molar excess of ATP exhibits the bent ATP conformation that is reminiscent of the class II tRNA synthetase family. The rotation of the sugar moiety about the glycosidic bond and the phosphate chain of the bent ligand in protein DX is even more pronounced, and the $\gamma$-phosphate of the bound molecule forms a close intramolecular polar contact with the 2′-hydroxyl of the sugar moiety. Unlike biological proteins, the ATP configuration in DX interacts with itself to maintain its strained conformation, with the conformation being facilitated

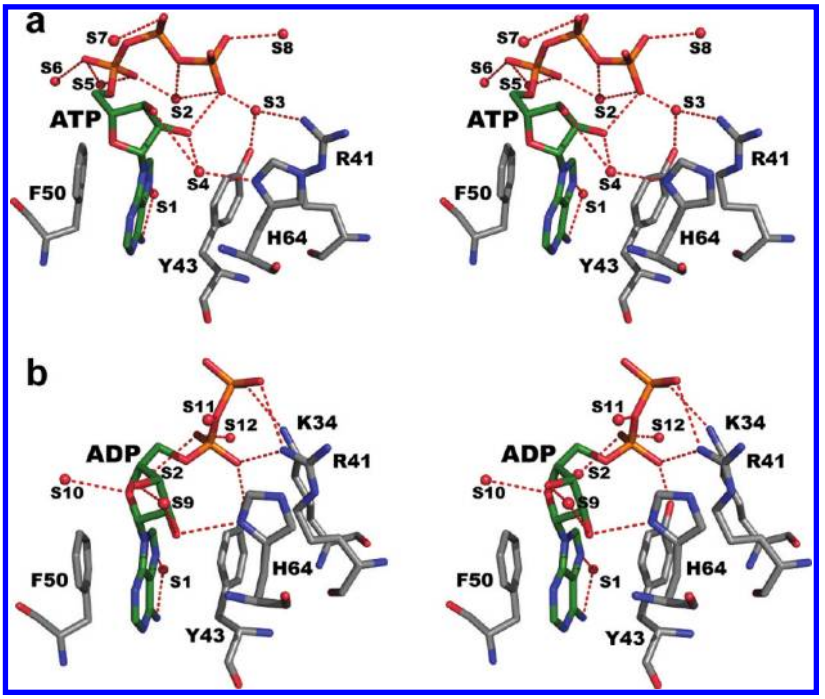


FIGURE 7: Stereoviews of the high-resolution DX structures crystallized in the presence of high and low concentrations of ATP. Three-dimensional structures of (a) DX grown in a 10-fold excess of ATP for 3 days and (b) DX grown under equimolar ATP conditions for 3 days, as reported previously (*11*). Hydrogen bonds shown as dotted lines were defined as any polar interaction between either solvent or side chain atoms with the bound ATP or ADP at a distance of ≤3.5 Å. The main chain carbonyl atoms of Met 45 and Gly 63 also contribute hydrogen bonds to the N6 and N3 atoms of the purine ring but have been omitted for the sake of clarity. Side chain atoms are colored by atom type (gray for carbon, red for oxygen, and blue for nitrogen). Bound ADP and ATP ligands are shown as stick models with colors by atom type (green for carbon, red for oxygen, blue for nitrogen, and orange for phosphate). Water numbers (signified by "S" to notate the red spheres) were selected on the basis of conserved positions between the two structures. Numbers that differ in position by a >1.0 Å radius in the superimposed models are noted with different solvent numbers. The corresponding solvent numbering with the PDB coordinate file waters is as follows: 1:10 DX-3 day, S1−S8 (96, 130, 155, 147, 144, 143, 160, and 159, respectively); 2P09, S1−S2 (510 and 549, respectively) and S9−S12 (535, 574, 571, and 542, respectively).

by only a small number of protein and solvent contacts, without the aid of any metal cofactors.

*Implications for Enzymatic Activity*. These results demonstrate that synthetic proteins can be evolved de novo from pools of random sequences that exhibit both ligand binding affinity and ligand conformations that are consistent with catalytic activity. The major implication of the results is that the first enzyme catalysts may have evolved from primordial pools of random sequences by adapting to bind substrates in conformations that closely approximate the Michaelis−Menten complex. The ability to trap a Michaelis complex in the crystalline state is consistent with the theory of substrate strain as originally proposed by Haldane (*33*) and later substantiated by Schultz and co-workers in their crystal structure of a strained porphyrin substrate bound to a catalytic antibody (*34*). Together, these structures led to the conclusion that these proteins bind ATP in a conformation that lowers the activation barrier for hydrolysis, leading to the presence of ADP in most of the structures.

This evidence of catalytic behavior is supported by a combination of biochemical data from solution binding assays, mass spectrometry, and thin-layer chromatography (*14*). High-resolution MALDI-TOF mass spectrometry of dissolved protein crystals shows a significant amount of ADP for DX crystallized in the presence of ATP. In contrast, only ATP is observed for Y43F crystallized under identical conditions. Both of these outcomes have been verified by thin-layer chromatography. The predominance of ATP in Y43F and ADP in DX is consistent with a 23- and 2.5-fold selectivity for ADP compared to ATP as measured in binding assays for Y43F and DX, respectively. Together, these experiments confirm the enzymatic reaction for DX in the crystalline state.

The mutations and ATP binding site are all located on the surface with a significant solvent exposure rather than buried inside the protein. While sequestration of ligands and critical amino acid residues is often considered to be critical for the control of their properties, small proteins such as DX and 18-19 have a relatively small volume to bury residues or ligands. Despite their presence on the surface, alterations of residues 32 and 65 significantly enhance the protein solubility and stability. The effect of these mutations is indirect; for example, the Asp 65 to Val alteration eliminates a salt bridge between Asp 65 and Arg 21, causing a significant repositioning of the Arg 21 side chain and a change in the electrostatic interactions involving Arg 21, Arg 31, and Asp 65. Although the ligand has an unusual exposure to the surface when compared to most biological proteins, the bound ATP/ADP ligand still possesses many of the interactions that are commonly found in natural proteins (*11*). This positioning on the surface has the advantage in that this relatively small protein can incorporate bound water molecules in stabilizing the bent configuration of the ligand. These observations suggest surface residues and ligands can play critical roles in the structure and function of proteins.

*Summary*. Several striking outcomes are observed in the six structures of the de novo proteins reported here. The overall structures of all of these proteins are very similar, with the same folds and backbones that superimpose closely; however, the bound cofactors have very different conformations. Despite the addition of ATP to the crystallization solutions in all cases, the D65V protein always has ADP bound, the Y43F protein always has ATP bound, and the DX structures have ADP or ATP bound depending upon the specific crystallization conditions.

Together, these results show the nucleobase is bound in a conventional manner through stacking interactions involving aromatic residues and hydrogen bonding to a bound water molecule, which is positioned by Tyr 43. In contrast to biological proteins, the phosphate atoms of the bound ATP are arranged in a highly distorted manner that is suggestive of a transition-state conformation that would reduce the energy required to convert ATP to ADP in a slow but catalytic process with multiple turnovers. An open question from these studies is how the observed ligand conformations influence the energetics. A variety of computational methods have been developed to calculate the free energies of ligands, but their application requires a balance between speed and accuracy. Because simpler approaches based upon empirical functions with fewer interactions result in faster but less reliable estimates (*35*), detailed calculations to understand how the structure can lower the ~8.5 kcal/mol activation barrier for ATP hydrolysis are planned and will be presented in a later publication.

This interpretation of the structures suggests that simple proteins are capable of metal-free catalysis and that design-free methods can be used to generate novel protein scaffolds with tailor-made functions. This work demonstrates that synthetic proteins have great potential to evolve with respect to both ligand binding affinity and protein stability. While the methodology has been applied only to ATP, the selection of proteins from the library can be extended to any other ligand. To select for catalytic activity, it is necessary to develop a selection strategy that separates active molecules from the inactive pool. While the field of de novo protein evolution is still in its infancy, techniques have been established for the evolution of ribozymes with specific, predefined functions from large pools of random sequences (*36*−*39*). Such studies have established a template for the experimental approaches that can be used to introduce catalytic function into a natural protein fold and create designer catalysts.

## ACKNOWLEDGMENT

Synchrotron data for this study were measured at beamline X12b of the National Synchrotron Light Source, which receives its principal financial support from the Offices of Biological and Environmental Research, Basic Energy Sciences of the U.S. Department of Energy, and from the National Center for Research Resources of the National Institutes of Health. We thank beamline scientists Drs. Dieter Schneider and Alexei Soares at NSLS X12b for their technical support during data collection.

## REFERENCES

1. Benner, S. A., and Sismour, A. M. (2005) Synthetic biology. *Nat. Rev. Genet. 6*, 533–543.
2. Endy, D. (2005) Foundations for engineering biology. *Nature 438*, 449–453.
3. Lippow, S. M., and Tidor, B. (2007) Progress in computational protein design. *Curr. Opin. Biotechnol. 18*, 305–311.
4. Kuhlman, B., Dantas, G., Ireton, G. C., Varani, G., Stoddard, B. L., and Baker, D. (2003) Design of a novel globular protein fold with atomic-level accuracy. *Science 302*, 1364–1368.
5. Bolon, D. N., and Mayo, S. L. (2001) Enzyme-like proteins by computational design. *Proc. Natl. Acad. Sci. U.S.A. 98*, 14274–14279.
6. Kaplan, J., and DeGrado, W. F. (2004) De novo design of catalytic proteins. *Proc. Natl. Acad. Sci. U.S.A. 101*, 11566–11570.
7. Rothlisberger, D., Khersonsky, O., Wollacott, A. M., Jiang, L., DeChancie, J., Betker, J., Gallaher, J. L., Althoff, E. A., Zanghellini, A., Dym, O., Albeck, S., Houk, K. N., Tawfik, D. S., and Baker, D. (2008) Kemp elimination catalysts by computational enzyme design. *Nature 453*, 190–195.
8. Watkins, J. L., and Chaput, J. C. (2008) Searching combinatorial libraries for native proteins with novel folds. *ChemBioChem 9*, 1361–1363.
9. Keefe, A. D., and Szostak, J. W. (2001) Functional proteins from a random-sequence library. *Nature 410*, 715–718.
10. Lo Surdo, P., Walsh, M. A., and Sollazzo, M. (2004) A novel ADP- and zinc-binding fold from function-directed in vitro evolution. *Nat. Struct. Mol. Biol. 11*, 382–383.
11. Smith, M. D., Rosenow, M. A., Wang, M., Allen, J. P., Szostak, J. W., and Chaput, J. C. (2007) Structural Insights into the Evolution of a Non-Biological Protein: Importance of Surface Residues in Protein Fold Optimization. *PLoS One 2*, e467.
12. Chaput, J. C., and Szostak, J. W. (2004) Evolutionary optimization of a nonbiological ATP binding protein for improved folding stability. *Chem. Biol. 11*, 865–874.
13. Mansy, S. S., Zhang, J., Kummerle, R., Nilsson, M., Chou, J. J., Szostak, J. W., and Chaput, J. C. (2007) Structure and Evolutionary Analysis of a Non-biological ATP-Binding Protein. *J. Mol. Biol. 371*, 501–513.
14. Simmons, C. R., Stomel, J. M., McConnell, M. D., Smith, D. A., Watkins, J. L., Allen, J. P., and Chaput, J. C. (2009) A synthetic protein selected for ligand binding affinity mediates ATP hydrolysis. *ACS Chem. Biol. 4*, 649–658.
15. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307–326.
16. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriaty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) PHENIX: Building new software for automated crystallographic structure determination. *Acta Crystallogr. D58*, 1948–1954.
17. Terwilliger, T. C. (2003) SOLVE and RESOLVE: Automate structure solution and density modification. *Methods Enzymol. 374*, 22–36.
18. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2126–2132.
19. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D50*, 760–763.
20. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D53*, 240–255.
21. Vaguine, A. A., Richelle, J., and Wodak, S. J. (1999) SFCHECK: A unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr. D55*, 191–205.
22. Laskowski, R. A., MacArthur, M. W., Moss, D. G., and Thornton, J. M. (1993) Procheck: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr. 26*, 283–291.
23. DeLano, W. L. (2002) The PYMOL molecular graphics system, DeLano Scientific, San Carlos, CA.
24. Stockwell, G. R., and Thornton, J. M. (2006) Conformational diversity of ligands bound to proteins. *J. Mol. Biol. 356*, 928–944.
25. Hulett, H. R. (1970) Non-enzymatic hydrolysis of adenosine phosphates. *Nature 225*, 1248–1249.
26. Schmitt, E., Moulinier, L., Fujiwara, S., Imanaka, T., Rhierry, J. C., and Moras, D. (1998) Crystal structure of aspartyl-tRNA synthetase from *Pyrococcus kodakaraensis* KOD: Archaeon specificity and catalytic mechanism of adenylate formation. *EMBO J. 17*, 5227–5237.
27. Kraft, L., Sprenger, G. A., and Lindqvist, Y. (2002) Conformational changes during the catalytic cycle of gluconate kinase as revealed by X-ray crystallography. *J. Mol. Biol. 318*, 1057–1069.
28. Traxler, P., and Furet, P. (1999) Strategies toward the design of novel and selective protein tyrosine kinase inhibitors. *Pharmacol. Ther. 82*, 195–206.
29. Denessiouk, K. A., Rantanen, V. V., and Johnson, M. S. (2001) Adenine recognition: A motif present in ATP-, CoA, NAD-, NADP-, and FAD-dependent proteins. *Proteins 44*, 282–291.
30. Pyrkov, T. V., Kosinsky, Y., Arseniev, A. S., Priestle, J. P., Jacoby, E., and Efremov, R. G. (2007) Complementarity of hydrophobic properties in ATP-protein binding: A new criterion to rank docking solutions. *Proteins 66*, 388–398.
31. Arnez, J. G., Augustine, J. G., Moras, D., and Francklyn, C. S. (1997) The first step of aminoacylation at the atomic level in histidyl-tRNA synthetase. *Proc. Natl. Acad. Sci. U.S.A. 94*, 7144–7149.
32. Eriksson, M., Uhlin, U., Ramaswamy, S., Ekberg, M., Regnstrom, K., Sjoberg, B. M., and Eklund, H. (1997) Binding of allosteric

effectors to ribonucleotide reductase protein R1: Reduction of active-site cysteines promotes substrate binding. *Structure 5*, 1077–1092.
33. Haldane, J. B. S. (1930) Enzymes, Longmans, London.
34. Yin, J., Andryski, S. E., Beuscher, A. E., Stevens, R. C., and Schulz, P. G. (2003) Structural evidence for substrate strain in antibody catalysis. *Proc. Natl. Acad. Sci. U.S.A. 100*, 856–861.
35. Singh, N., and Warshel, A. (2010) Absolute binding free energy calculations: On the accuracy of computational scoring of protein-ligand interactions. *Proteins 78*, 1705–1723.
36. Wilson, D. S., and Szostak, J. W. (1999) *In vitro* selection of functional nucleic acids. *Annu. Rev. Biochem. 68*, 611–647.
37. Doudna, J. A., and Cech, T. R. (2002) The chemical repertoire of natural ribozymes. *Nature 418*, 222–228.
38. Joyce, G. F. (2002) The antiquity of RNA-based evolution. *Nature 418*, 214–221.
39. Seelig, B., and Szostak, J. W. (2007) Selection and evolution of enzymes from a partially randomized non-catalytic scaffold. *Nature 448*, 828–831.