PAS (PER-ARNT-SIM) domains are a family of sensor protein domains involved in signal transduction in a wide range of organisms. Recent structural studies have revealed that these domains contain a structurally conserved α/β-fold, whereas almost no conservation is observed at the amino acid sequence level. The photoactive yellow protein, a bacterial light sensor, has been proposed as the PAS structural prototype yet contains an N-terminal helix-turn-helix motif not found in other PAS domains. Here we describe the atomic resolution structure of a photoactive yellow protein deletion mutant lacking this motif, revealing that the PAS domain is indeed able to fold independently and is not affected by the removal of these residues. Computer simulations of currently known PAS domain structures reveal that these domains are not only structurally conserved but are also similar in their conformational flexibilities. The observed motions point to a possible common mechanism for communicating ligand binding/activation to downstream transducer proteins.

PAS1 domains are structural modules that can be found in proteins in all kingdoms of life (1, 2). The PAS module was first identified in the Drosophila clock protein PER and the basic helix-loop-helix containing transcription factors ARNT (aryl-hydrocarbon receptor nuclear translocator) in mammals and SIM (single-minded protein) insects (3). Most PAS domains are sensory modules, typically sensing oxygen tension, redox potential, or light intensity (1, 4). Alternatively, they mediate protein-protein interactions or bind small ligands (5). Although the amino acid sequences of the different PAS domains show little similarity, their three-dimensional structures appear to be conserved. All of the PAS domains resemble the structure of photoactive yellow protein (PYP) (4), a photoreceptor presumed to be involved in a phototactic response of the bacterium Ectotherhodospira halophila to intense blue light (6). Its structure reveals an α/β-fold with the light-sensitive chromophore p-coumaric acid bound to the protein via a thioester linkage (7). It is the only PAS domain of which the catalytic function, i.e. signal generation and transduction, has been studied in great detail. The protein has been shown to undergo a photocycle linked to isomerization of the chromophore (8–12). The ground state (pG) has a UV-visible absorbance maximum at 446 nm. After absorption of a blue photon, the protein returns from the primary excited state into the first transient ground state, a strongly red-shifted intermediate, at the picosecond time scale (13–17). A more moderately red-shifted intermediate absorbing maximally at 465 nm is formed on the nanosecond time scale (18). The red-shifted intermediate spontaneously converts into a blue-shifted intermediate absorbing maximally at 355 nm at the sub-millisecond time scale (18, 19). The blue-shifted intermediate subsequently relaxes back to pG on a sub-second time scale (15, 18–20) or faster in a light-dependent reaction (21, 22). Several detailed studies, including Laue diffraction and cryo-crystallography (9, 11, 12), NMR spectroscopy (23), small angle x-ray scattering (24, 25), biochemical experiments (26), and Fourier transform infrared spectroscopy (27, 28) and computer simulations (29, 30), have revealed that during the PYP photocycle distinct significant conformational changes occur. It is these conformational changes that are thought to translate the photon signal into a cellular response via subsequent protein/protein interactions. To study the possible protein motions involved in the signal photocycle, PYP dynamics have been investigated by computer simulation (29). This study suggested that chromophore-linked concerted motions may be present in pG and that these motions might be amplified upon isomerization of the chromophore. The simulations, later supported by x-ray crystallographic studies (30), also suggest that conserved glycines were serving as hinge points, allowing substructures in the protein to fluctuate relative to each other. In a subsequent study where the rigidity of the PYP backbone was altered by mutation of these glycines, the role of these hinge points in the signal transduction process was further confirmed (31). The glycines that were investigated in this study fall within the PAS-fold (4) and show a large degree of conservation throughout the PAS family. This has led to the speculation that apart from a conserved structure the PAS domains may have similar conformational freedom and associated signal transduction mechanism (30).

Here we investigate whether the dynamic properties of PAS domains are intrinsic properties associated with their conserved fold. First, we have mutated the PYP from E. halophila into a minimal PAS domain by the removal of the N-terminal cap (see also Ref. 32). To be able to tackle the dynamic properties of this minimal PAS domain, its three-dimensional struc-
structure was refined against 1.14-Å synchrotron diffraction data. Second, this structure was used in a comparative computational study on the conformational flexibility of all of the PAS domains for which crystals structures are available: HERG, the N-terminal domain of a human potassium channel (33); LOV2, a photoreceptor domain from plants (34); and FixL, a bacterial oxygen sensor (35). Essential dynamics analyses on the sampled configurational space of all of these PAS domains reveal conserved concerted motions. This supports the hypothesis that the common structure of PAS domains implies common flexibility and that it is this conserved property that is fundamental for PAS domain function in signal transduction.

MATERIALS AND METHODS

Crystallographic, Diffraction, and Refinement

Δ25PYP encompassing residues 26–125 of PYP was expressed and purified as described previously (32). Crystals were grown by equilibration of 1 μl of 30 mg/ml protein with 1 μl of mother liquor (1.8 M ammonium sulfate, 10 mM CoCl₂, 100 mM MES, pH 6.5) against a 1-ml reservoir of mother liquor. Crystals appeared after 2–3 days with a largest dimension of 0.4 mm.

Diffraction data were collected at beamline ID14-EH1 (European Synchrotron Radiation Facility, Grenoble, France) and processed with the HKL package (Table I) (36). The structure of Δ25PYP was solved by molecular replacement with AMoRe (37) using the native PYP structure of turkey lysozyme (Protein Data Bank code 135L) bearing no structural resemblance to PAS domains was also calculated. During the refinement. cryo-cooled to 100 K. All measured data were included in structure refinement.

Table 1. Details of data collection and structure refinement

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution range (Å)</td>
<td>17.1-1.14 (1.18-1.14)</td>
</tr>
<tr>
<td>No. observed reflections</td>
<td>299780 (23225)</td>
</tr>
<tr>
<td>No. unique reflections</td>
<td>75208 (7010)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>4.0 (3.3)</td>
</tr>
<tr>
<td>I/σ(I)</td>
<td>13.5 (4.3)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>94.4 (89.2)</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.068 (0.217)</td>
</tr>
<tr>
<td>R cryst, R free</td>
<td>0.147, 0.177</td>
</tr>
<tr>
<td>No. groups</td>
<td>200 residues, 406 H2O</td>
</tr>
<tr>
<td>Root mean square deviation from ideal geometry</td>
<td>0.010</td>
</tr>
<tr>
<td>Bonds (Å)</td>
<td>2.0</td>
</tr>
<tr>
<td>Angles (°)</td>
<td>2.6</td>
</tr>
<tr>
<td>B-factor root mean square deviation (Å²)</td>
<td>15.8 (protein), 33.5 (water)</td>
</tr>
<tr>
<td>all bonds</td>
<td>2.6</td>
</tr>
</tbody>
</table>

Fig. 1. Structure and maps of Δ25PYP. Ribbon representation of the crystal structure of Δ25PYP. The asymmetric unit cell contains two proteins. Secondary structure elements are marked on the structures. The Fobs – Fcalc map just before including the chromophore is shown in magenta, contoured at 2.5 σ. Hydrophobic residues that have become solvent-exposed because of the deletion of residues 1–25 are shown as green sticks.

Fig. 2. Conformational changes. Positional shifts of equivalent Cα atoms after superposition of the two Δ25PYP monomers in the asymmetric unit on wtPYP and on each other.
CONCOORD runs, 1000 structures were generated and a damping factor of 0.25 was applied to avoid unreasonable side chain geometries. Essential Dynamics—Essential dynamics (43) determines concerted motions of atoms from an ensemble of structures, for example, a set of crystal structures (44–47) or a trajectory from a computer simulation (43, 48–51). Here the CONCOORD ensembles were used as input. A covariance matrix is constructed that describes the correlation of the positional shifts of one atom with those of another atom as shown in Equation 1,

\[
C_{ij} = \langle (x_i - x_{i,0})(x_j - x_{j,0}) \rangle \tag{Eq. 1}
\]

where \(x_i\) and \(x_j\) represent the coordinates of atoms \(i\) and \(j\) in a conformation, whereas \(x_{i,0}\) and \(x_{j,0}\) represent the average coordinates of the atoms over the ensemble. The average is calculated over all structures after they are superimposed on a reference structure to remove overall translational and rotational motion. Diagonalizing this matrix yields a set of eigenvectors and eigenvalues. The eigenvectors are directions in a 3N-dimensional space (where \(N\) is the number of atoms), and motion along a single eigenvector corresponds to concerted displacements of groups of atoms in Cartesian space. The eigenvalues are a measure of the mean square fluctuation of the system along the corresponding eigenvectors. The eigenvectors are sorted according to their eigenvalue, the first eigenvector having the largest eigenvalue.

To allow direct comparison of concerted motions for different proteins, an equal number of atoms must be used in the essential dynamics calculations. A first simplification is that only Cα atoms are taken into account, which sufficiently represent the large motions of the protein backbone (48, 52). When the structures also contain insertions and deletions such as in the PAS domains (Fig. 3), further simplifications will need to be applied to reduce all of the structures to a common core (44). Residues in the PAS domains that overlapped structurally were selected by the DALI server (53), which performs a pairwise comparison of secondary structure elements. The results of the pairwise alignment on secondary structure were compared to yield the common structural elements present in the PAS domains (Fig. 3). For lysozyme, the negative control, an equal number of residues was selected starting from the N terminus.

RESULTS AND DISCUSSION

Atomic Resolution Crystal Structure of \(\Delta_{25}\)PYP—The structure of \(\Delta_{25}\)PYP was solved by molecular replacement and refined to a 1.14-Å resolution (R-factor = 0.147, \(R_{free} = 0.177\))

Fig. 3. Comparison of the PAS cores. Ribbon representations of the structures of wtPYP (7), the \(\Delta_{25}\)PYP mutant described here, FixL (35), HERG (33), LOV2 (34), and turkey lysozyme (Protein Data Bank code 135L). Residues that align structurally (see also Fig. 4) and are used for comparisons are highlighted in red. The cofactors, if present, are drawn in a yellow colored stick model. Homologous residues at similar structural positions (Fig. 4) are depicted in green.
(Fig. 1 and Table I). The asymmetric unit contains the protein molecules related by a non-crystallographic 2-fold rotation axis (Fig. 1). The molecules have a similar conformation with a root mean square deviation of 0.77 Å on Cα atoms. Compared with the wtPYP structure, the two molecules superimpose with root mean square deviations of 0.99 and 0.76 Å, respectively. From these superpositions, positional shifts of the Cα atoms of the mutant structures with respect to the positions of the Cα atoms in wild type PYP are given in Fig. 2. The N terminus and the loops consisting of residues 84–88, 98–101, and 111–117 in Δ25PYP have a different conformation than those in wild type PYP. The different conformation of the Δ25PYP N terminus compared with the equivalent residues in wtPYP is most probably caused by the deletion of the first 25 residues (Fig. 3). When the first two residues at the N terminus of wtPYP are excluded from the superposition, the root mean square deviation is reduced by ~0.2 Å. From the NMR structure and the comparison of two crystal forms of wild type PYP, the loop around residue Met-100 is observed to be flexible (7, 30, 54). Close contacts between the two monomers in the asymmetric unit cell affect the conformation of the “100 loop” (Fig. 1). The distance between the backbone atoms of the two Met-100 residues is <4.0 Å and could influence the conformation of this loop. The large differences in the orientation of the 111–117 loop, which is disordered in this loop and previously reported PYP structures (30, 31, 41), and the 84–88 loop can again be explained by crystal contacts (Fig. 1).

Native PYP contains two hydrophobic cores, one within the PAS domain between the β-sheet and the αC helix and another between the β-sheet and the two small helices of the N-terminal domain (7). For the latter domain, residues Phe-28 (at the start of βA), Trp-199, and Phe-121 (both on βE) extend toward the N-terminal domain and have become solvent exposed in Δ25PYP yet appear not to self-dimerize through crystal contacts. This finding agrees with the observation that in Δ25PYP the fluorescence emission of Trp-199 (the only tryptophan in PYP) is enhanced and blue-shifted, suggesting a more polar environment (32). It is likely that it is these solvent-exposed hydrophobic residues that cause the reported decrease in temperature stability observed for Δ25PYP (32). In summary, the structural data show that the removal of the first 25 residues does not significantly affect the overall fold of the PAS domain core. This is in agreement with the spectrophotometric data on Δ25PYP, which show that the Δ25PYP absorbance maximum associated with the chemical environment of the chromophore and very sensitive to perturbation is only minimally blue-shifted (32). Similar photocycle intermediates as in wtPYP are present. Only the kinetics of the recovery reaction in the photocycle is slowed down, which again could be associated with a less stable ground state structure due to exposure of several hydrophobic residues.

As observed earlier (4, 34), none of the other currently known PAS domain structures (Fig. 3) possesses extra residues similar to the N terminus of wtPYP. Thus, this may be a unique feature that plays a specific but, as yet, unidentified role in PYP function. Although the PAS domains share a common fold, only few residues are conserved at the sequence level (Figs. 3 and 4). When these residues (mostly leucines, isoleucines, and valines) are compared among the different PAS domains, they appear to form part of a conserved hydrophobic core (Figs. 3 and 4). Interestingly, most of these residues are located in a cluster near helix αA, which appears to act as a lid on the ligand binding pockets and undergoes conformational changes in PYP (9, 11, 23, 25, 29, 30).

### Table II

<table>
<thead>
<tr>
<th>Protein</th>
<th>PYP</th>
<th>FixL</th>
<th>HERG</th>
<th>LOV2</th>
<th>Δ25PYP</th>
<th>Lysozyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>PYP</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FixL</td>
<td>0.70</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HERG</td>
<td>0.66</td>
<td>0.72</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOV2</td>
<td>0.69</td>
<td>0.71</td>
<td>0.73</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δ25PYP</td>
<td>0.78</td>
<td>0.68</td>
<td>0.69</td>
<td>0.70</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Lysozyme</td>
<td>0.24</td>
<td>0.24</td>
<td>0.22</td>
<td>0.24</td>
<td>0.23</td>
<td>1.00</td>
</tr>
</tbody>
</table>

**PAS Domain Flexibility**—Because PAS domains function as sensor proteins in signaling pathways and share a common fold, it is possible that they also have common dynamic properties, which would allow them to communicate with transducer proteins through a conserved mechanism. We have attempted to investigate this through computer simulation of all of the structurally characterized PAS domains, namely PYP (7), FixL (35), HERG (33), and LOV (34). The complete proteins were subjected to CONCOORD simulations (42) followed by extraction of the common Cα atoms as defined by a structure-based sequence alignment (Figs. 3 and 4). This encompassed 78 residues (marked in Figs. 3 and 4) including most of the central β-sheet, the αA/β helices, and part of the long αC helix. The resulting ensembles of structures were analyzed by essential dynamics (43), yielding separate sets of eigenvectors that describe concerted fluctuations of atoms for each protein. These are sorted by their corresponding eigenvalues, i.e., the first...
eigenvector being the one with the largest eigenvalue, revealing in all cases that the majority (>95%) of the motion is covered by the first 5% (12) of the eigenvectors. With this condensed description of flexibility in the individual PAS domains, comparisons are facilitated.

Sets of eigenvectors can be projected onto each other yielding a cumulative square inner product, indicating the degree of similarity of the motions described by the eigenvectors. Here we have focused on the first 12 eigenvectors (5% total eigenvectors), because these together describe approximately 95% of the total motion in the ensembles. Table II shows that the eigenvectors from the different PAS domains are very similar, suggesting that the cores of the PAS domains share common motions, which are not present in lysozyme (the negative control). This is further confirmed by projection of the PAS domain eigenvectors onto the first three eigenvectors calculated from the wtPYP ensemble (Fig. 5). Whereas the other PAS domains reproduce these largest wtPYP motions for up to 90% within the first 12 eigenvectors, they are almost absent in the lysozyme ensemble. Thus, the PAS domains not only share a common structure but also share a common conformational flexibility.

To understand the motions described by the eigenvectors on a molecular level, the minimum and maximum projections onto an eigenvector can be translated back to Cartesian space and compared as C traces. In Fig. 6, the minimum and maximum projections of the first 3 eigenvectors of Δ25PYP are compared. The central β-sheet appears to be relatively static, whereas the loops, most notably the αA/αB segment, show the largest fluctuations. In the PAS domains, this segment is generally important for the binding of the ligand (7, 34, 35). For instance, in PYP, residue Arg-52 on this segment is known to undergo a conformational change (9, 11, 29, 30) upon isomerization of the chromophore. Glu-46, which shares a proton with the chromophore, is also located in this region (Fig. 4). Similarly, the αA/αB segment is involved in binding the heme in FixL (35) and the FMN in LOV (34) both via interaction with a phenylalanine, which lies at the equivalent position of Glu-46 in PYP. In addition, a recent analysis of LOV domains has revealed that the αA/αB region participates in a conserved salt bridge, which is also observed in FixL and HERG and has been proposed to be involved in signal transduction (55). It is noteworthy that despite these similar interactions and conservation of conformational flexibility, there is almost no sequence conservation in the αA/αB segment.

CONCLUSION

The data presented here show that in the absence of the N-terminal domain, PYP maintains its PAS-fold despite the exposure of several hydrophobic residues to solvent. The Δ25PYP structure together with the recently determined of the LOV domain in complex with FMN (34) allowed further structural comparisons of the PAS family. Although these proteins have almost entirely dissimilar sequences, their structures are remarkably similar with the conserved parts, the
β-sheet and the αA/B helices, making up the PAS core. This finding suggests that although these proteins bind different ligands, their signaling states are reached through similar conformational changes. We investigated this by simulating the complete PAS domain proteins that have been structurally defined to date and extracting from that the structurally conserved core. An analysis of the data shows that in particular the αA/B segment moves in a concerted fashion. Thus, we propose that despite the absence of any sequence conservation, the PAS domains are not only structurally conserved but also share a common conformational flexibility that may have evolved to (i) accommodate the various input signals from different ligands/co-factors located at different positions in the domain and (ii) transmit the sensing event to downstream transducer proteins.

Acknowledgments—We thank Raafat Hamze for contributions in the early stages of the project. We thank the European Synchrotron Radiation Facility (Grenoble, France) for the time at beamline ID14-EH1.

REFERENCES
PROTEIN STRUCTURE AND FOLDING:
PAS Domains: COMMON STRUCTURE
AND COMMON FLEXIBILITY

Jocelyne Vreede, Michael A. van der Horst,
Klaas J. Hellingwerf, Wim Crielaard and
Daan M. F. van Aalten
doi: 10.1074/jbc.M301701200 originally published online March 14, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M301701200

Find articles, minireviews, Reflections and Classics on similar topics on the JBC Affinity Sites.

Alerts:
  • When this article is cited
  • When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 57 references, 13 of which can be accessed free at
http://www.jbc.org/content/278/20/18434.full.html#ref-list-1