NMR structure of the apoptosis- and inflammation-related NALP1 pyrin domain

Sebastian Hiller[1,4], Andreas Kohl[2,4], Francesco Fiorito[1], Torsten Herrmann[1], Gerhard Wider[1], Jürg Tschopp[3], Markus G. Grütter[2] and Kurt Wüthrich[1]

[1]*Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule Zürich, CH-8093 Zürich, Switzerland*

[2]*Biochemisches Institut der Universität Zürich, CH-8057 Zürich, Switzerland*

[3]*Institut de Biochimie, Université de Lausanne, CH-1066 Epalinges, Switzerland*

[4]*These authors contributed equally to this work.*

Correspondance:
Markus G. Grütter
+41-1-635 5581 (phone)
+41-1-635 68 34 (fax)
gruetter@bioc.unizh.ch

Summary

Signalling in apoptosis and inflammation is often mediated by proteins of the death domain superfamily in the Fas/FADD/Caspase-8 or the Apaf-1/Caspase-9 pathways. This superfamily currently comprises the death domain (DD), death effector domain (DED), caspase recruitment domain (CARD), and pyrin domain (PYD) subfamilies . The PYD subfamily is most abundant, but three-dimensional structures are only available for the subfamilies DD, DED and CARD, which have an antiparallel arrangement of six α-helices as common fold. This paper presents the NMR structure of PYD of NALP1, a protein that is involved in the innate immune response and is a component of the inflammasome. The structure of NALP1 PYD differs from all other known death domain superfamily structures in that the third α-helix is replaced by a flexibly disordered loop. This unique feature appears to relate to the molecular basis of familial Mediterranean fever (FMF), a genetic disease caused by single point mutations.

Introduction

Apoptosis and inflammation are two closely related processes in the cells of higher organisms. Signal transduction and protein oligomerization in both apoptosis and inflammation are often mediated by a group of protein–protein interaction domains, the death domain superfamily (Hengartner, 2000; Weber et al., 2001). The death domain superfamily currently comprises four subfamilies, the death domain (DD), the death effector domain (DED), the caspase recruitment domain (CARD) and the pyrin domain (PYD or PAAD, DAPIN, PLM, PYRN, PyK) (Bertin et al., 2000; Martinon et al., 2001; Staub et al., 2001). After its discovery in the protein pyrin (or marenostrin) (The French FMF Consortium, 1997; The International FMF Consortium, 1997), PYD was recognized to be a fourth member of the death domain superfamily, based on sequence alignments, secondary structure predictions and homology modelling (Bertin et al., 2000; Martinon et al., 2001; Staub et al., 2001). So far, only protein–protein interactions between members of the same DD subfamily have been observed, i.e., PYD–PYD, DD–DD, CARD–CARD and DED–DED interactions, indicating high specificity of the binding mechanisms between death domain superfamily members (Weber et al., 2001). Examples are the Fas/FADD/Caspase-8, the TNF-R/RAIDD/RIP or the Apaf-1/Caspase-9 pathways. Even though PYD is the most abundant subfamily of all the death domain superfamily members, structural information is only available for the DD, DED and CARD domains (Huang et al., 1996; Eberstadt et al., 1998; Zhou et al., 1999; Berglund et al., 2000). All reported structures of the death domain superfamily share an antiparallel arrangement of six α-helices as a common fold. The lengths and the relative orientations of the six helices vary between the individual structures, but the overall topology, termed the death domain fold, is conserved.

NALP1 (DEFCAP, CARD7, NAC) (Hlaing et al., 2001; Chu et al., 2001) is a multidomain protein involved in inflammation and apoptosis. It is composed of a N-terminal PYD, a NACHT (domain present in neuronal apoptosis inhibitory protein (NAIP), the major histocompatibility complex (MHC), class II transactivator (CIITA), HET-E, TP1) (Koonin et al., 2000), a NAD (NACHT associated domain), a LRR (leucine rich repeat) and a CARD domain at the C-terminus (Tschopp et al., 2003). This multidomain architecture is shared by many other proteins (e.g., Apaf-1) and they are now classified as CATERPILLAR proteins (Harton et al., 2002). NALP1 interacts with ASC (apoptosis-associated speck-like protein containing a CARD) (Martinon et al., 2001; Masumoto et al., 1999) via PYD, and with procaspase-5 via the C-terminal CARD. These proteins, together with procaspase-1 form the "inflammasome", a 700 kDa multiprotein-complex (Martinon et al., 2002). The stoichiometry of the proteins in the inflammasome is not known, but the inflammasome has been shown to activate procaspase-1, which ultimately processes the potent endogenous pyrogen IL-1β from its precursor pro-IL-1β (Martinon et al., 2002). Single-point mutations in CATERPILLER proteins can have severe consequences. Examples are the development of autoinflammatory diseases, such as familial cold autoinflammatory syndrome (FCAS) or Muckle-Wells syndrome (MWS) upon mutations in cryopyrin (NALP3, PYPAF1) (Kastner et al., 2001), and familial Mediterranean fever (FMF) upon mutations in pyrin (The French FMF Consortium, 1997; The International FMF Consortium, 1997). One of these FMF associated point mutations, R42W, is located in the PYD domain of pyrin (Hull et al., 2003).

We have been able to express, solubilize and purify the full length PYD domain of NALP1, and in this paper we present its NMR structure. The structure differs

significantly from the death domain superfamily fold, as the third α-helix is replaced by a flexibly disordered loop. This difference is of high importance, especially because the third helix can play a critical role in protein–protein interactions within the death domain superfamily (Weber et al., 2001; Kaufmann et al., 2002). A sequence alignment of several PYD domains reveals further structural features relevant for the protein–protein interactions, and allows locating the FMF associated point mutation in the PYD of pyrin in the flexible loop replacing the third α-helix.

Results and Discussion

Protein expression and purification

As an isolated domain, NALP1 PYD aggregates at concentrations above 10 μM. Therefore we used a two-domain fusion protein of the full-length NALP1 PYD and a 56-residue G B1 solubility enhancing tag (Zhou et al., 2001), which could be concentrated in the mM range for NMR experiments (see Experimental Procedures).

The NMR structure of NALP1 PYD

The NMR structure of NALP1 PYD (Table 1, Fig. 1 and Fig. 2) consists of a tightly packed arrangement of five α-helices (Fig. 1 b), connected by three short linkers of non-regular secondary structure, and a long flexibly disordered polypeptide segment connecting the second and third α-helices. The five helices comprising the residues 8−13, 18−30, 51−59, 63−76 and 80−85 are numbered 1, 2, 4, 5 and 6 to maintain consistency with the 6-helix death domain superfamily fold. The surface of NALP1 PYD shows a large negatively charged patch consisting of E19 and E20 (Fig. 1 c), which turns out to be a common trait of PYD domains (see below).

Evidence for the flexibly disordered loop

For the polypeptide segment between the helices 2 and 4, no long-range NOEs to protons in the remaining part of the domain were observed. This is reflected in the low precision of the structure determination in this region (Fig. 1 a). The flexible nature of the structural disorder was confirmed by measurements of heteronuclear $^{15}N\{^{1}H\}$–NOEs (Fig. 3 a), which showed that the effective rotational correlation time for residues

in this loop is significantly shorter than in the other parts of the protein. The ^{13}C chemical shift deviations for C^{α} and C^{β} atoms from random coil values (Fig. 3 b) show large positive values for the five α-helices, whereas the linkers between the helices and especially the loop between α_2 and α_4 show only small deviations, indicating non-regular secondary structure. In this context, it is noteworthy that the T41–P42 peptide bond appears to undergo cis-trans isomerization (Wüthrich et al., 1984): In the [^{15}N,^{1}H]-COSY spectrum the residues flanking P42 show resonance signals for a second species, with an abundance of about 10%.

The PYD subfamily

In the genome of *Homo sapiens*, 18 PYDs have been described so far (Harton et al., 2002; Tschopp et al., 2003), and for 12 of them, i.e., NALP 1, 2, 4, 5, 6, 7, 11 and 12, cryopyrin, pyrin, ASC and ASCII, the amino acid sequence has been published in the SWISS-PROT database. These sequences can be aligned for a high degree of homology (Fig. 4 a), whereby four residues are strictly conserved in all 12 PYDs. These are the aforementioned P42 in the loop linking α_2 and α_4, and the hydrophobic residues L13, L16 and F24, which are part of the protein core. Residues 9, 21, 28, 53, 57, 74 and 77 are conserved hydrophobic residues in all 12 PYDs, which are also part of the hydrophobic core of the domain (Fig. 4 a), suggesting that a similar fold to NALP1 PYD can be expected for the other PYDs as well. Further conserved features are located on the surface of the domain, where two acidic and one basic regions are present in all 12 PYDs. These include E14, which is conserved in all PYDs except NALP5, which instead contains E15 (Fig. 4 a). Further, residue 20 is acidic in all sequences except NALP6, where residue 19 is acidic. In 7 out of the 12 sequences, the residues 19 and 20

are both acidic, which includes NALP1 PYD (Fig. 1 c). The conserved basic patch is a cluster of the four sidechains of residues 22, 23, 25 and 47, which are closely spaced in NALP1 PYD. In all 12 PYDs, these four residues have polar sidechains. Residue 25 is basic in all but one of the proteins, in each sequence the cluster carries at least one positive charge, and in eight of the PYDs there are 3 or 4 positive charges. Similar alignments of DDs, DEDs and CARDs showed no comparable degree of conservation within each of these families (data not shown).

Pyrin PYD and the molecular basis of FMF

As mentioned above, the amino acid exchange R42W in the PYD-domain of pyrin presents with FMF, a disease that causes ongoing episodes of inflammatory attacks to the patient. Residue R42 of the pyrin PYD-domain corresponds to residue A43 in the structure of NALP1 PYD. The location of this amino acid in the flexible loop, next to the conserved P42 and the conserved basic surface patch (yellow in Fig. 4 a), suggests that this part of the domain is involved in protein–protein interactions that trigger apoptosis or inflammation. The conserved acidic regions might further promote contacts via intermolecular saltbridges with the basic patch of another PYD. It seems even conceivable that the loop between the helices 2 and 4 might adopt a helical structure upon complex formation. In another scenario, the loop 3 could be structured and become α-helical by an intra-protein domain-domain interaction which might play a role in the assembly or regulation of NALP1. However no evidence for any of these possibilities was found. Our experiments show no secondary structure and no contacts between residues in the loop 3 and residues in the rest of the PYD (see evidence for the flexibly disordered loop).

Comparison of the four death domain subfamilies

A structural comparison between members of the four subfamilies of the death domain superfamily (Fig. 4 b) shows similarities and differences. Common to all four structures is the spatial arrangement of the α-helices 1, 2, 4, 5 and 6. They have slightly different tilt angles with respect to each other in the four structures, but the overall topology is the same. A striking difference is the absence of α-helix 3 in NALP1 PYD, which makes it unique in the death domain superfamily. This variation has also been observed in a mutant form of the Fas DD, the "proliferation mutant", where the single-point mutation V238N causes the α-helix 3 to unfold. Val 238, which is located in the loop connecting the α-helices 2 and 3, forms hydrophobic interactions with Val 233 of α_2, Ile 243 of α_3, and Leu 268 and Trp 265 of α_4. Replacement of Val 238 by Asn disrupts these hydrophobic interactions, resulting in the unfolding of α-helix 3 (Eberstadt et al., 1997). Mice carrying this variant gene present with lymphadenopathy and suffer from a systemic autoimmune disease (Watanabe-Fukunaga et al., 1992). However, all wild type structures of DDs, DEDs or CARDs reported so far include the α-helix 3. At this stage, it is not clear whether the flexible loop is a general feature of all PYDs. When compared to the other death domain subfamilies, NALP1 PYD lacks hydrophobic residues in the loop polypeptide segment linking α_2 and α_4, which is not true for other PYDs. Whether the presence of some hydrophobic sidechains would be sufficient to induce a regular secondary structure in the place of the flexible loop cannot be decided based on the available data. Secondary structure predictions applied to the PYD consensus sequence have come to contradictory results with regard to the

polypeptide segment corresponding to the α-helix 3 in the other death domains (Martinon et al., 2001; Staub et al., 2001).

Conclusions

In conclusion, the three-dimensional structure of a PYD-domain shows that the fold of NALP1 PYD is locally different from the known death domain superfamily folds in the region between the α-helices 2 and 4. This observation motivates further studies, to investigate whether the absence of α-helix 3 is a general feature of all PYDs. If this were the case, PYD should be considered to form a new subgroup of the death domain superfamily, with the DD, DED and CARD domains all belonging to another subgroup. The structure of NALP1 PYD allows insight into possible interactions between multiple PYDs, and provides a molecular model visualizing mutation sites that present with genetic fever diseases. It could thus also become a target for drug discovery projects in the area of inflammation.

Experimental Procedures

Cloning and Protein expression

The NALP1 PYD gene (aa 1–92) was subcloned from the full length Nalp1 and inserted into a pET20b+ (Novagen) derived vector, which contained an N-terminal G B1 expression and solubility enhancer with a C-terminal 6-His tag (Zhou et al., 2001). For the NMR experiments, ^{13}C, ^{15}N uniformly labelled protein was obtained by using the BL21(DE3) (Novagen) *E. coli* expression strain and a standard M9 minimal medium prepared with $^{15}NH_4Cl$ and with [$^{13}C_6$]-glucose. The degree of isotope labelling of the protein after purification was confirmed to be >95% using ESI-MS.

Sample preparation

The protein was purified using a Ni-NTA affinity column (Quiagen) followed by a size exclusion column (sephadex 75, 26/60, Pharmacia). The NMR sample used for the structure determination was prepared at pH 6.5 in 95% H_2O / 5% D_2O containing 50 mM Na_2HPO_4, 50 mM NaCl, 1 mM CHAPS, 20 mM [$^2H_{10}$]-DTT (Martek), 0.02 % NaN_3, 0.1 mM EDTA and a protease inhibitor cocktail (Roche, Complete). The concentration of NALP1 PYD was approximately 1 mM.

NMR data collection

All NMR data recorded for the structure determination were performed on the same sample at 20° C on Bruker DRX 600 or DRX 750 spectrometers. For the resonance assignment the following standard experiments were recorded: 3D HNCA, 3D HNCACB, 3D constant-time-HNCA, 3D CBCA(CO)NH, 3D HC(C)H-TOCSY, 3D HC(C)H-COSY (Wider, 1998). Distance constraints were obtained from three NOESY

experiments with mixing times of 60 ms, *i.e.*, 3D ^{15}N-resolved [^{1}H,^{1}H]-NOESY, and 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY with the ^{13}C carrier frequency in the aliphatic and aromatic region, respectively. The data sets used for obtaining the sequence-specific resonance assignments were interactively peak-picked using the programs XEASY (Bartels et al., 1995) and CARA (in-house software, unpublished). Measurements of heteronuclear NOEs were performed at 20° C on a Bruker DRX 500 spectrometer equipped with a cryoprobe. The NOE spectrum and the reference spectrum were integrated using CARA. Error bars have been calculated from the spectral noise and the combination of two independently recorded experiments. For Fig. 3 b, a 1-2-1 weighted three-point smoothing function was applied to the raw data for the presentation of the C^{α}/C^{β} chemical shift deviations from the random coil values.

Structure calculation

The aforementioned three NOESY spectra were picked and assigned with the automated NOESY peak picking and NOE assignment method ATNOS/CANDID (Herrmann et al., 2002a; 2002b) in the program DYANA (Güntert et al., 1997). The input for the iterative ATNOS/CANDID approach consisted of the chemical shift lists obtained from the sequence-specific resonance assignment, and the three NOESY spectra. ATNOS performs multiple cycles of NOE peak identification in conjunction with automated NOE assignment by the CANDID algorithm, and in the present implementation, structure calculation with the program DYANA. At the outset of the NOESY spectral analysis, the NOESY peak picking and NOE identification algorithm ATNOS makes use of the chemical shifts available from the independently obtained sequence-specific resonance assignment, and in more advanced stages of the calculation

also of the intermediate three-dimensional protein structure. Seven cycles of automated NOESY peak picking and NOE identification with ATNOS, automated NOE assignment with CANDID and subsequent structure calculation with the torsion angle dynamics algorithm of DYANA were performed. During the first six ATNOS/CANDID cycles, ambiguous distance constraints were used. 158 dihedral angle constraints for the backbone angles ϕ and ψ derived from C^{α} chemical shifts were added to the input for each cycle of the structure calculation. For the final structure calculation in ATNOS/CANDID cycle 7, a total of 4436 NOESY cross peaks were retained that could be unambiguously rationalized with the three-dimensional protein structure obtained in the ATNOS/CANDID cycle 6, leading to a total of 2088 meaningful NOE upper distance limits. The 20 conformers from the ATNOS/CANDID cycle 7 with the lowest final DYANA target function values were energy-minimized in a water shell with the program OPALp (Luginbühl et al., 1996). The program MOLMOL (Koradi et al., 1996) was used to analyze the resulting 20 energy-minimized conformers and to prepare the figures showing molecular models.

Use of structure data bases and sequence alignment

The amino acid sequences of the 12 PYDs in Fig. 4 a have been taken from the SWISS-PROT database. The sequence alignment was done using ClustalW (Higgins et al., 1994). Coordinates for the domains Apaf-1 CARD (Zhou et al., 1999), Fadd DD (Berglund et al., 2000) and Fadd DED (Eberstadt et al., 1998) have been taken from the PDB Protein Data Bank.

References

Bartels, C., Xia, T. H., Billeter, M., Güntert, P. and Wüthrich, K. (1995). The program XEASY for computer-supported NMR spectral-analysis of biological macromolecules. J. Biomol. NMR *6*, 1–10.

Berglund, H., Olerenshaw, D., Sankar, A., Federwisch, M., McDonald, N.Q. and Driscoll, P.C. (2000). The three-dimensional solution structure and dynamic properties of the human FADD death domain. J. Mol. Biol. *302*, 171–188.

Bertin, J. and DiStefano, P. S. (2000). The PYRIN domain: a novel motif found in apoptosis and inflammation proteins. Cell Death Differ. *7*, 1273–1274.

Chu, Z. L., Pio, F., Xie, Z., Welsh, K., Krajewska, M., Krajewski, S., Godzik, A. and Reed, J.C. (2001). A novel enhancer of the Apaf1 apoptosome involved in cytochrome c-dependent caspase activation and apoptosis. J. Biol. Chem. *276*, 9239–9245.

Eberstadt, M., Huang, B. H., Olejniczak, E. T. and Fesik, S. W. (1997). The lymphoproliferation mutation in Fas locally unfolds the Fas death domain. Nature Struct. Biol. *4*, 983–985.

Eberstadt, M., Huang, B., Chen, Z., Meadows, R.P., Ng, S.C., Zheng, L., Lenardo, M.J. and Fesik, S.W. (1998). NMR structure and mutagenesis of the FADD (Mort1) death-effector domain. Nature *392*, 941–945.

Güntert, P., Mumenthaler, C. and Wüthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. J. Mol. Biol. *273*, 283–298.

Harton, J. A., Linhoff, M. W., Zhang, J. and Ting, J. P.-Y. (2002). Cutting edge: CATERPILLER: A large family of mammalian genes containing CARD, Pyrin, nucleotide-binding, and leucine-rich repeat domains. J. Immunol. *169*, 4088–4093.

Hengartner, M. O. (2000). The biochemistry of apoptosis. Nature *407*, 770–776.

Herrmann, T., Güntert, P. and Wüthrich, K. (2002a). Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J. Mol. Biol. *319*, 209–227.

Herrmann, T., Güntert, P. and Wüthrich, K. (2002b). Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. J. Biomol. NMR *24*, 171–189.

Higgins, D., Higgins, D.G. and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. *22*, 4673–4680.

Hlaing, T., Guo, R.F., Dilley, K.A., Loussia, J.M., Morrish, T.A., Shi, M.M., Vincenz, C. and Ward, P.A. (2001). Molecular cloning and characterization of DEFCAP-L and -S, two isoforms of a novel member of the mammalian Ced-4 family of apoptosis proteins. J. Biol. Chem. *276*, 9230–9238.

Huang, B., Eberstadt, M., Olejniczak, E.T., Meadows, R.P. and Fesik, S.W. (1996). NMR structure and mutagenesis of the Fas (APO-1/CD95) death domain. Nature *384*, 638–641.

Hull, K. M., Shoham, N., Chae, J. J., Aksentijevich, I. and Kastner, D. L. (2003). The expanding spectrum of systemic autoinflammatory disorders and their rheumatic manifestations. Curr. Opin. Rheumatol. *15*, 61–69.

Kastner, D. L, and O'Shea J. J. (2001). A fever gene comes in from the cold. Nature Genet. *29*, 241–242.

Kaufmann, M., Bozic, D., Briand, C., Bodmer, J.L., Zerbe, O., Kohl, A., Tschopp, J. and Grutter, M.G. (2002). Identification of a basic surface area of the FADD death effector domain critical for apoptotic signaling. FEBS Lett. *527*, 250–254.

Koonin, E.V. and Aravind, L. (2000). The NACHT family – a new group of predicted NTPases implicated in apoptosis and MHC transcription activation. Trends in Biochem. Sci. *25*, 223–224.

Koradi, R., Billeter, M. and Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. J. Mol. Graph. *14*, 51–55.

Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R. and Thornton, J. M. (1996). AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR *8*, 477–486.

Luginbühl, P., Güntert, P., Billeter, M. and Wüthrich, K. (1996). The new program OPAL for molecular dynamics simulations and energy refinements of biological macromolecules. J. Biomol. NMR *8*, 136–146.

Martinon, F., Hofmann, K. and Tschopp, J. (2001). The pyrin domain: a possible member of the death domain-fold family implicated in apoptosis and inflammation. Curr. Biol. *11*, R118–R120.

Martinon, F., Burns, K. and Tschopp, J. (2002). The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proIL-1β. Mol. Cell *10*, 417–426.

Masumoto, J., Taniguchi, S., Ayukawa, K., Sarvotham, H., Kishino, T., Niikawa, N., Hidaka, E., Katsuyama, T., Higuchi, T. and Sagara, J. (1999). ASC, a novel 22-kDa protein, aggregates during apoptosis of human promyelocytic leukemia HL-60 cells. J. Biol. Chem. *274*, 33835–33838.

Staub, E., Dahl, E. and Rosenthal, A. (2001). The DAPIN family: a novel domain links apoptotic and interferon response proteins. Trends Biochem. Sci. *26*, 83–85.

The French FMF Consortium (1997). A candidate gene for familial Mediterranean fever. Nature Genet. *17*, 25–31.

The International FMF Consortium (1997). Ancient missense mutations in a new member of the RoRet gene family are likely to cause familial Mediterranean fever. Cell *90*, 797–807.

Tschopp, J., Martinon, F. and Burns, K. (2003). NALPS: A novel protein family involved in inflammation. Nature Rev, Mol. Cell Biol. *4*, 95–104.

Watanabe-Fukunaga, R., Brannan, C. I., Copeland, N. G., Jenkins, N. A. and Nagata, S. (1992). Lymphoproliferation disorder in mice explained by defects in Fas antigen that mediates apoptosis. Nature *356*, 314–317.

Weber, C. H. and Vincenz, C. (2001). The death domain superfamily: a tale of two interfaces? Trends Biochem. Sci. *26*, 475–481.

Wider, G. (1998). Technical aspects of NMR spectroscopy with biological macromolecules and studies of hydration in solution. Progr. NMR Spectrosc. *32*, 193–275.

Wüthrich, K., Billeter, M. and Braun W. (1984). Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton–proton distances. J. Mol. Biol. *180*, 715–740.

Zhou, P., Chou, J., Olea, R. S., Yuan, J. Y. and Wagner, G. (1999). Solution structure of Apaf-1 CARD and its interaction with caspase-9 CARD: A structural basis for specific adaptor/caspase interaction. Proc. Natl. Acad. Sci. USA *96*, 11265–11270.

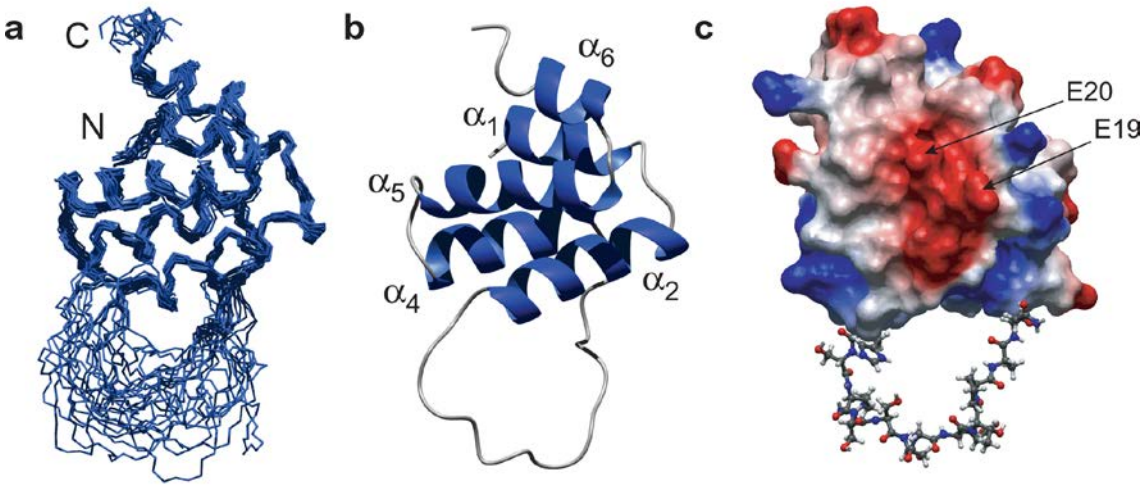
Zhou, P., Lugovskoy, A. A. and Wagner, G. (2001). A solubility-enhancement tag (SET) for NMR studies of poorly behaving proteins. J. Biomol. NMR *20*, 11–14.

Acknowledgements

We thank Gerhard Wagner (Harvard Medical school) for the GB1 gene, Oliver Zerbe (ETH Zürich) for initial NMR measurements, the Swiss National Science Foundation and the Swiss National Center of Competence in Research in Structural Biology for financial support.

Correspondence

Should be addressed to M. G. G. (gruetter@bioc.unizh.ch). Atomic coordinates for a bundle of 20 conformers of NALP1 PYD have been deposited in the PDB Protein Data Bank under the accession code 1PN5.



gruetter_fig1

Figure 1 NMR structure of the NALP1 PYD-domain (residues 6 to 89 are shown). **a**, The polypeptide backbone represented by a bundle of 20 energy-minimized DYANA (Güntert et al., 1997) conformers. The chain ends are identified by the letters N and C. **b**, Ribbon drawing of one of the 20 energy-minimized conformers. α_1, α_2 and α_4–α_6 identify the five α-helices. **c**, Surface view of the same conformer as in b. Relative to panels a and b, the structure has been rotated by 45 degrees around a vertical axis, and is shown in the same orientation as in Fig. 4 a. Acidic regions of the surface are painted red, basic regions are painted blue. Arrows indicate the locations of the residues E19 and E20.

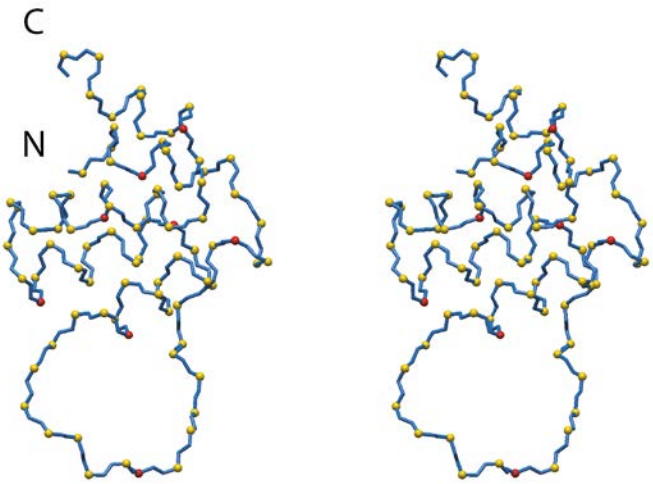


Figure 2 Stereoview of the same conformer as in Fig. 1b (residues 6 to 89 are shown). The polypeptide backbone is drawn in blue. C^{α} atoms with the residue number being a multiple of 10 are represented by red balls, all other C^{α} atoms by yellow balls. The chain ends are identified by the letters N and C.

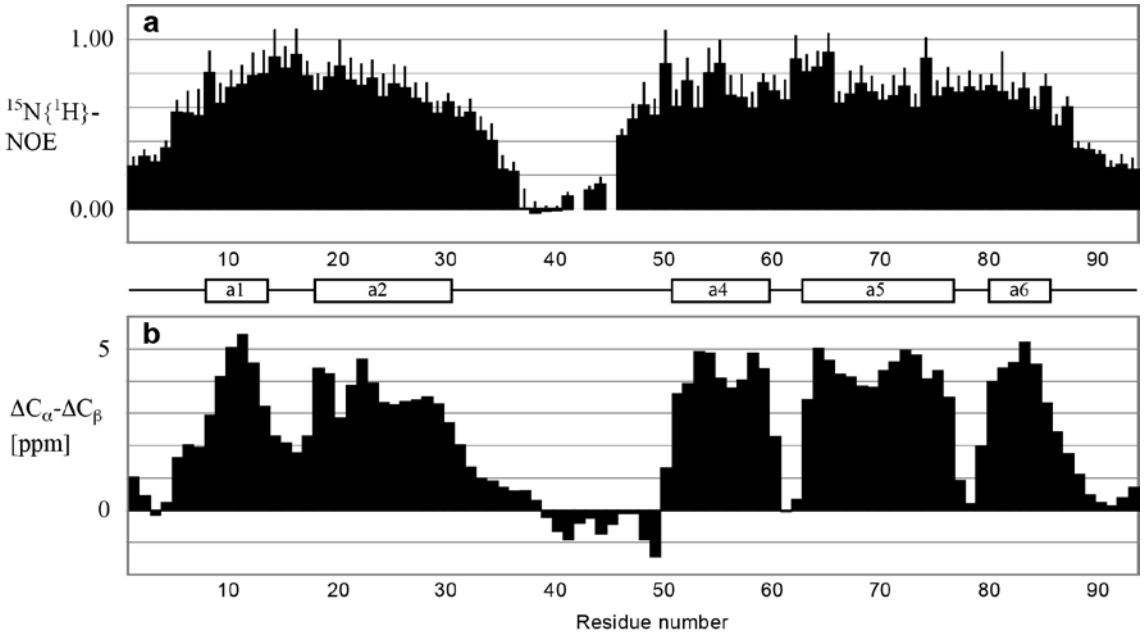


Figure 3 Experiments showing that NALP1 PYD is flexibly unstructured between helices $\alpha 2$ and $\alpha 4$. **a**, Heteronuclear $^{15}N\{^{1}H\}$-NOE measurements. Values between 0.5 and 1 indicate rigid parts of the protein, values <0.5 manifest increased flexibility. **b**, Differences between the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts measured for NALP1 PYD and the corresponding random coil values. Values >2.5 are typical for α-helical residues. The locations of regular secondary structures in NALP1 PYD are shown below panel a.

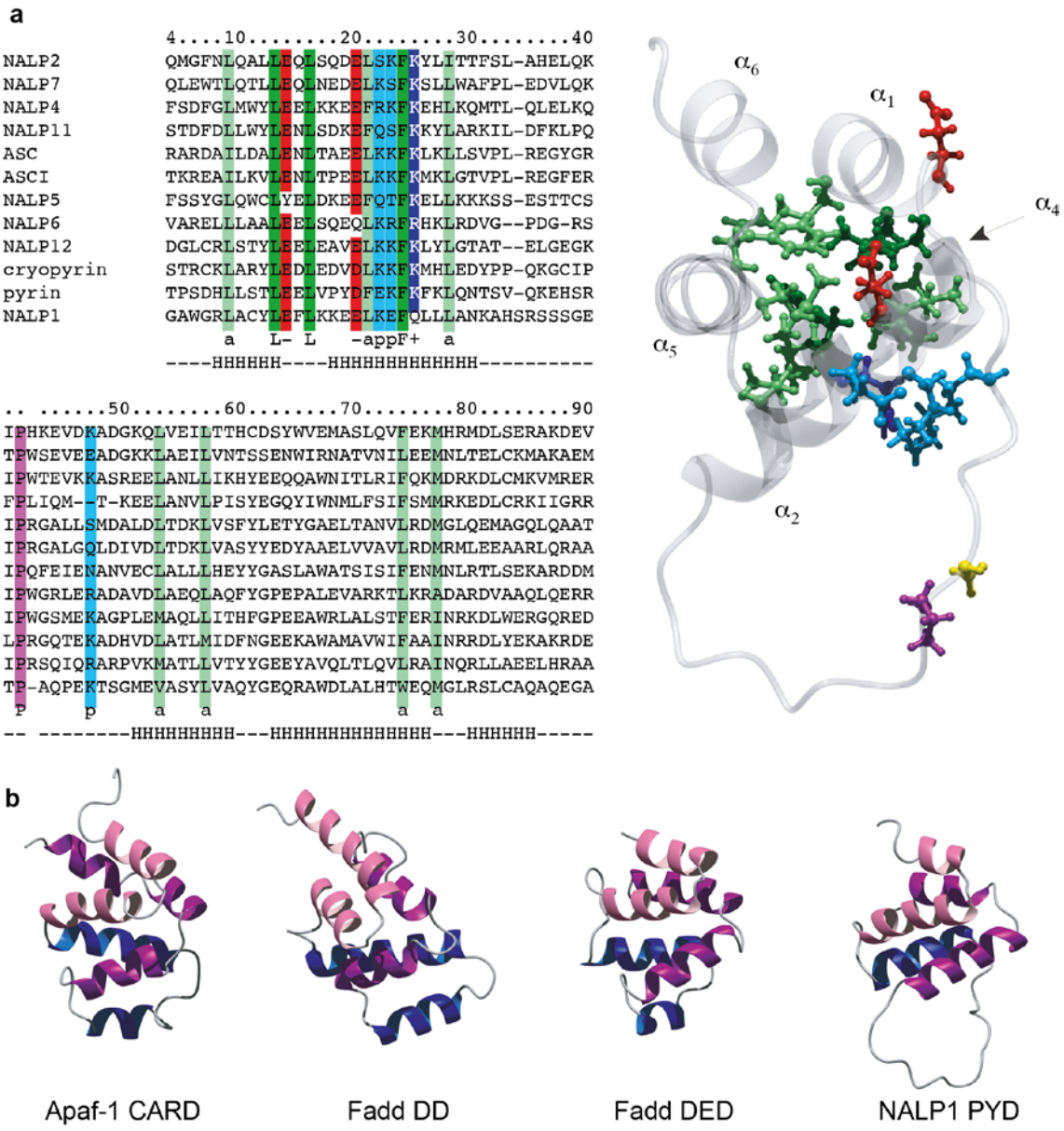


Figure 4 Structural analysis of the PYD family. **a**, Sequence alignment of the PYD domains from 12 *Homo sapiens* proteins. The residue numbering of NALP1 PYD is given at the top. Common features of the 12 PYDs are identified by colouring and by a marker below the sequences: Strictly identical residues (dark green or violet, one-letter code of the conserved amino acid), conservation of polar residues (light blue, p), hydrophobic residues (light green, a), residues 14 and 20 (red, –), residue 25 (dark blue, +). The bottom line indicates the location of the α-helices (H) in the structure of NALP1 PYD. On the right, the conserved residues are visualized in the structure of NALP1 PYD,

using the same colour code as in the sequences. In addition, residue A43 is displayed in yellow (see text). **b**, Folds of the death domain superfamily. For each of the four subfamilies, a member is shown in the ribbon representation: Apaf-1 CARD (Zhou et al., 1999), Fadd DD (Berglund et al., 2000), Fadd DED (Eberstadt et al., 1998) and NALP1 PYD. The helices have been coloured as follows: α_1 and α_2 violet, α_3 and α_4 blue, α_5 and α_6 pink.

Table 1 Input for the structure calculation and characterization of the energy-minimized NMR structure of NALP1 PYD

Quantity	Value
NOE upper distance limits	2088
Dihedral angle constraints	158
Residual DYANA target function value (Å^2)* †	1.79 ± 0.66
Residual NOE distance limit violations*	
Number ≥ 0.1 Å	32 ± 5
Maximum (Å)	0.15 ± 0.01
AMBER energies, kcal/mol*	
Total	-5963.66 ± 163.61
van der Waals	-423.91± 28.88
Electrostatic	-6852.28 ± 146.62
rmsd from ideal geometry*	
Bond lengths, (Å)	0.0074 ± 0.0002
Bond angles (degrees)	1.95 ± 0.06
rmsd from the mean coordinates (Å)‡	
bb G B1 (all 56 residues)	0.58 ± 0.11
ha G B1 (all 56 residues)	1.11 ± 0.14
bb NALP1 PYD (7–32, 47–87)	0.55 ± 0.08
ha NALP1 PYD (7–32, 47–87)	1.20 ± 0.14

Ramachandran plot statistics§	
Most favored regions (%)	76
Additional allowed regions(%)	20
Generously allowed regions (%)	3
Disallowed regions (%)	1

*The numbers given are the average value ± the standard deviation for the set of 20 energy-minimized conformers used to represent the NMR structure of the complete construct of G B1 + NALP1 PYD.

†Before energy minimization.

‡The rmsd values (average ± standard deviation for the bundle of 20 conformers) were calculated separately for the G B1 domain and for NALP1 PYD. bb indicates the backbone atoms N, C^{α}, C', and ha stands for all heavy atoms; the numbers in parentheses indicate the residues for which the rmsd was calculated.

§ As determined by PROCHECK (Laskowski et al., 1996)