



NMR structural studies of the first catalytic half-domain of ubiquitin activating enzyme



Mariusz Jaremko^{a,1}, Łukasz Jaremko^{a,b,1}, Michał Nowakowski^c, Marek Wojciechowski^d, Roman H. Szczepanowski^d, Renata Panecka^d, Igor Zhukov^{e,f}, Matthias Bochtler^{d,e}, Andrzej Ejchart^{e,*}

^a Max Planck Institute for Biophysical Chemistry, Department for NMR-based Structural Biology, Am Fassberg 11, 37077 Göttingen, Germany

^b Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE), Am Fassberg 11, 37077 Göttingen, Germany

^c Faculty of Chemistry, Biological and Chemical Research Centre, University of Warsaw, Żwirki i Wigury 101, 02-089 Warsaw, Poland

^d International Institute of Molecular and Cell Biology, Trojdena 4, 02-109 Warsaw, Poland

^e Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Pawińskiego 5A, 02-106 Warsaw, Poland

^f NanoBioMedical Centre, Adam Mickiewicz University, Umultowska 85, 61-614 Poznan, Poland

ARTICLE INFO

Article history:

Received 13 August 2013

Received in revised form 30 October 2013

Accepted 31 October 2013

Available online 6 November 2013

Keywords:

Ubiquitin-activating enzyme E1

First catalytic cysteine half-domain (FCCH)

Second catalytic cysteine half-domain

(SCCH)

NMR structure determination

¹⁵N nuclear magnetic relaxation

NMR study interactions in solution

Protein structure

Backbone dynamics

ABSTRACT

We report a high resolution NMR structure and ¹⁵N relaxation studies of the first catalytic cysteine half-domain (FCCH) of the mouse ubiquitin-activating enzyme E1, together with interaction studies of FCCH and the other catalytic E1 subdomain – SCCH (second catalytic cysteine half-domain). In solution, mouse FCCH forms a well-defined six-stranded antiparallel β-barrel structure, a common fold for many proteins with a variety of cellular functions. ¹⁵N relaxation data reveal FCCH complex backbone dynamics and indicate which residues experience slow intramolecular motions. Some of these residues make contacts with the polar face of ubiquitin in the co-crystal structure of yeast E1 and ubiquitin. However, the titration of FCCH with ubiquitin does not show any visible chemical shift changes in the 2D ¹H/¹⁵N HSQC spectra of the FCCH. The 2D ¹H/¹⁵N HSQC experiments performed both for each catalytic half-domain individually and for their equimolar mixture in the millimolar concentration range display no detectable chemical shift perturbation, suggesting a lack of interaction between the two subdomains unless they are covalently linked via the adenylation domain.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Ubiquitin is a major posttranslational protein modifier in all eukaryotes. The highly conserved 76 amino acid protein can be isopeptide linked via its C-terminal glycine residue to the ε-amino group of a lysine residue in a substrate protein. If the substrate protein is ubiquitin itself, polyubiquitin chains are formed, which serve as a proteasome targeting signal. Ubiquitylation requires the subsequent action of three enzymes: E1 (ubiquitin activating

enzyme), E2 (ubiquitin conjugating enzyme), and E3 (ubiquitin ligase) that together transfer Ub to substrate proteins (Hershko and Ciechanover, 1998; Cooper, 2000).

The ubiquitin activating enzyme E1 consumes ATP and converts Ub to a transfer-competent, enzyme-bound thioester. The reaction begins with Ub-adenylate formation and the release of pyrophosphate. The active site cysteine of the E1 then displaces the AMP, leading to a ubiquitin-E1 thioester complex. More detailed studies have shown that ubiquitin activation follows a complex choreography involving concerted processing of two Ub molecules (Haas et al., 1982). In recent years the structure of yeast E1 with a Ub molecule bound in the adenylation site has been reported (Lee and Schindelin, 2008). The E1 enzyme has several domains: an adenylation domain (composed of active and inactive adenylation subdomains), a so-called catalytic cysteine domain, and smaller accessory domains: a four helix bundle and a ubiquitin fold domain. The catalytic cysteine domain itself is composed of subdomains, which have been termed the first (FCCH) and second (SCCH) catalytic cysteine half domains (Szczepanowski et al., 2005) (Fig. 1).

Abbreviations: AMP, adenosine-5'-monophosphate; ATP, adenosine-5'-triphosphate; CPMG, Carr–Purcell–Meiboom–Gill pulse train; E1, ubiquitin activating enzyme; E2, ubiquitin conjugating enzyme; E3, ubiquitin ligase; FCCH, first catalytic cysteine half-domain; HSQC, heteronuclear single quantum correlation; NOESY, nuclear Overhauser effect spectroscopy; RMSD, root mean square deviation; SCCH, second catalytic cysteine half-domain; TRIS-d₁₁, perdeuterated 2-amino-2(hydroxymethyl)1,3-propanediol; Ub, ubiquitin.

* Corresponding author.

E-mail address: aejchart@ibb.waw.pl (A. Ejchart).

¹ These authors contributed equally to this work.

Both the FCCH (111 amino acid residues) and the SCCH (276 amino acid residues) fold autonomously in solution. Previously it was shown that the SCCH of mouse E1 forms crystals suitable for X-ray diffraction experiments (Szczepanowski et al., 2005). Despite the publication of the yeast E1 – ubiquitin complex crystal structure, the detailed catalytic mechanism of ubiquitylation initiation is still not well understood.

In this study, we present an NMR determined structure of the autonomously folded FCCH subdomain of mouse E1, coupled with ^{15}N backbone relaxation studies. Moreover, we report that NMR cannot detect interactions between the FCCH and ubiquitin, or between FCCH and SCCH if they are on separate polypeptide chains.

2. Materials and methods

2.1. Protein expression and purification

The Swiss-Prot entry Q02053 for mouse E1 has recently been changed by the addition of 60 amino acids to the amino terminus of the protein, reflecting a reassignment of the initiator methionine. For consistency with previous structural work on mouse E1 (Szczepanowski et al., 2005), we based our numbering on the earlier version of the entry throughout the manuscript and also in the PDB submission. The design of the *Escherichia coli* expression constructs for the mouse E1 FCCH (residues 202–312 of the full length sequence according to Swiss-Prot entry Q02053) and SCCH of mouse E1 (residues 626–891) with a histidine tag has been described previously (Szczepanowski et al., 2005). The protocol available from the EMBL peptide services (EMBL Heidelberg) was used for the expression of ^{15}N and $^{13}\text{C}/^{15}\text{N}$ labeled proteins. 5 ml of medium A (Supplementary Table A.1) was inoculated with a single colony of *E. coli* BL21 (DE3) and grown overnight at 37 °C. This overnight culture was added to 1.1 of medium A and grown to an OD_{600} of 0.7–1.0. Afterwards, the culture was shifted to 28 °C. Expression was induced by adding IPTG to a concentration of

0.5 mM. Cells were grown for additional 4 h and harvested by centrifugation. Recombinant, labeled proteins were purified by IMAC chromatography, followed by gel filtration (Szczepanowski et al., 2005).

2.2. NMR assignment and structure determination

^{13}C , ^{15}N -double labeled protein was suspended in a 90 : 10 $\text{H}_2\text{O}/\text{D}_2\text{O}$ mixture, 50 mM TRIS- d_{11} , 50 mM NaCl at pH 6.5 (uncorrected value) at a 1 mM concentration. 650 μl aliquots were used for NMR experiments. All NMR experiments used for structure determination were performed at 11.7 T on a Varian Unity + 500 spectrometer at 298 K. FCCH and SCCH interactions in solution were studied by $^1\text{H}/^{15}\text{N}$ HSQC measurements of either ^{15}N -labeled FCCH subdomain only or a 1:1 mixture of ^{15}N labeled FCCH with unlabeled SCCH. Sample preparation, buffer composition and measurement conditions were the same as for the double labeled protein. FCCH titration with ubiquitin was carried out on 0.6 mM ^{13}C , ^{15}N -double labeled FCCH, using the buffer and other conditions as described above. Molar ratios of FCCH to ubiquitin were varied between 0 and 5.

NMR data were processed in NMRPipe (Delaglio et al., 1995) and analyzed with Sparky (Goddard and Kneller, 2010) and CARA software (Keller, 2004). Sequence specific assignment of backbone and side chain ^1H , ^{13}C and ^{15}N resonances was based on standard 3D techniques and described previously (Jaremko et al., 2006). Briefly, the ^1H , ^{15}N , and ^{13}C backbone resonances were assigned using 3D HNCACB (Wittekind and Mueller, 1993), 3D HNCA (Ikura et al., 1990), and HNCO (Muhandiram and Kay, 1994) spectra and confirmed by the ^{15}N -edited 3D NOESY-HSQC (mixing time 150 ms) spectrum (Zhang et al., 1994). Assignments of aliphatic ^1H and ^{13}C resonances in side chains were done based on 3D HBHA(CO)NH (Grzesiek and Bax, 1993), C(CO)NH-TOCSY (Gardner et al., 1996), and HCCH-TOCSY (Bax et al., 1990) experiments.

NOESY distance constraints were derived from 3D heteronuclear ^{15}N - and ^{13}C -edited NOESY-HSQC experiments (Muhandiram et al., 1993). The initial structure calculations were performed with CYANA software (Guentert et al., 1997). The automatic NOESY assignment procedure provided distance constraints. Additional restraints for backbone φ and ψ torsion angles were defined based on N, H^{N} , $\text{C}\alpha$, $\text{C}\beta$, $\text{H}\alpha$ and C' chemical shifts estimated using the PREDITOR server (Berjanskii et al., 2006). Additional hydrogen bond constraints, defined as $r_{\text{HN-O}} = 1.7 \dots 2.0 \text{ \AA}$ and $r_{\text{N-O}} = 2.7 \dots 3.0 \text{ \AA}$ were introduced based on geometric criteria before the final structure calculations. A hydrogen bond was selected for the final refinement in explicit solvent if it existed in more than 75% of structures in the ensemble. XPLOR-NIH 2.26 (Schwieters et al., 2003) was used for final structure calculations on the 70 lowest energy structures from 200 submitted for simulated annealing. The edge length of the water box was 18.8 \AA . The final ensemble discussed in this article represents 20 of the lowest energy conformers without distance violations of more than 0.5 \AA and dihedral angle violations of more than 5° obtained after water refinement. Evaluation of the FCCH structure quality was performed with PROCHECK-NMR (Laskowski et al., 1996) and WHAT-IF (Vriend, 1990) programs.

2.3. ^{15}N relaxation data measurements

All ^{15}N relaxation measurements were performed with a ^{15}N -labeled protein sample on a Varian 400 MHz (9.4 T) spectrometer at 298 K. Longitudinal (R_1) and transverse (R_2) relaxation rates were measured using a sensitivity-enhanced ^{15}N - ^1H HSQC pulse sequence (Kay et al., 1992) included in the ProteinPack Varian Inc. (Palo Alto, USA) software. Zero filling was performed prior to the Fourier transformation. Eight evolution periods (10, 60, 110, 170, 240, 330, 460, and 600 ms) were used for the determination

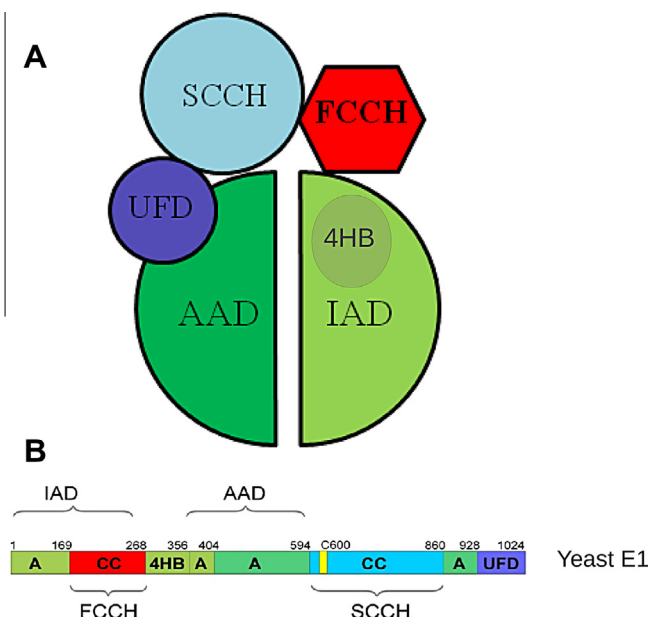


Fig. 1. (A) Schematic architecture of E1 enzyme. The enzyme consists of an adenylation domain, a cysteine catalytic domain (CC), a four helix bundle (4HB), and a ubiquitin fold domain (UFD). The adenylation domain can be further subdivided into inactive (IAD) and active (AAD) adenylation (sub)domains, whereas the catalytic cysteine domain is split into first (FCCH) and second (SCCH) catalytic cysteine subdomains. (B) Subunit and domain architecture in the yeast E1. The adenylation domains (A) are built from discontinuous regions of the amino acid sequence. The position of the catalytic cysteine residue has been indicated. Panel B is closely modeled on an earlier figure by Lee and Schindelin (2008).

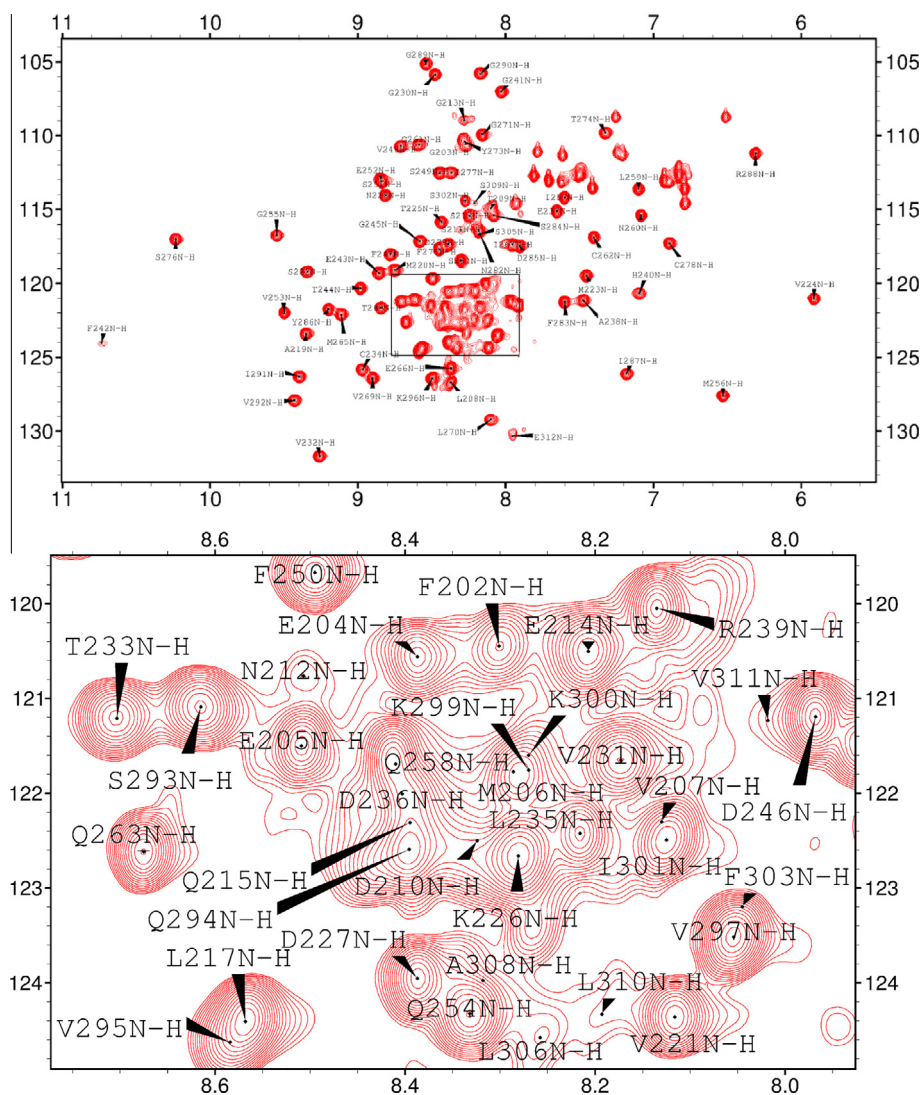


Fig. 2. $^1\text{H}/^{15}\text{N}$ HSQC spectrum of the FCCH domain of mouse activator protein E1 measured at 11.7 T. Full range of amide correlations – upper part of the figure. The frame in the central part marks the enlarged area shown in the lower part of the figure. The terminal NH_2 groups of Asn and Gln residues are not labeled.

of R_1 values. R_2 relaxation rate measurements were performed with the Carr–Purcell–Meiboom–Gill (CPMG) pulse train. Refocusing time during eight evolution periods (10, 30, 50, 70, 90, 110, 130, and 170 ms) was 650 μs . In both types of experiments, 3.0 s delays between the scans were used. $\{^1\text{H}\}$ - ^{15}N NOEs (Kay et al., 1989) were measured with the pulse sequence included in the ProteinPack Varian Inc. (Palo Alto, USA) software. NOE values were calculated from spectra measured with and without ^1H presaturation (3.0 s). The recycling delay was 8.0 s. Resonance intensities were used for calculating relaxation rates and NOE values. Experimental errors of relaxation rates were obtained from appropriate elements of the variance–covariance matrix. Experimental errors of NOE values were evaluated from signal-to-noise ratios obtained for corresponding signals in spectra with and without NOEs (Fushman, 2003).

2.4. Analysis of ^{15}N relaxation data

Relaxation data were analyzed using the model-free approach (Lipari and Szabo, 1982). The number of global parameters, a , describing overall diffusion tumbling depends on the relevant model of motion; $a = 1$ for the isotropic diffusion, $a = 4$ for the axially symmetric diffusion, and $a = 6$ for the anisotropic diffusion.

(Tjandra et al., 1995; Woessner, 1962). Three local, residue-specific parameters comprise a generalized order parameter S^2 , which is a measure of the degree of spatial restriction of the motions faster than the overall diffusion, an effective correlation time τ_{int} related to the rate of these motions, and R_{ex} describing conformational exchange contribution to the transverse relaxation rates resulting from the dynamic processes slower than the overall rotational diffusion, but fast on the chemical shift time scale. These processes are most often characterized by the microsecond to millisecond time scale (Stone et al., 1992). The R_{ex} contribution to the transverse relaxation rate is proportional to the square of the chemical shift difference between exchanging states, $\Delta\delta$, and to ω_N , the Larmor frequency. It should be pointed out that the conformational exchange mechanism can affect the apparent transverse relaxation rate only if $\Delta\delta \neq 0$. The optimal model parameters were determined by the least squares procedure consisting of minimization through a grid search of the target function comprising the sum of the squared differences between the experimental values of the relaxation parameters and their model-derived counterparts (Stone et al., 1992; Nowakowski et al., 2011, 2013). Model parameter uncertainties derived in the minimization of target function were obtained as standard deviations from 200 Monte Carlo simulations (Press et al., 1986).

For the analysis of the relaxation data measured at a single magnetic field strength, the unfavorable observations to parameter ratio was taken into account. In order to reduce the number of parameters, the R_1R_2 product as a function of amino acid sequence was used to separate residues exhibiting chemical exchange (group A) from those with $R_{ex} = 0$ (group B). Initially, residues with the R_1R_2 product exceeding the weighted mean value of R_1R_2 more than three standard deviations are assigned to group A. The use of R_1R_2 product rather than their quotient (R_2/R_1) made it possible to minimize the effect of expected motional anisotropy (Kneller et al., 2002). After the preliminary selection of residues based on the R_1R_2 values the minimization procedure was performed. Residues assigned to group A, for which $R_{ex} = 0$, were moved to the group B. Conversely, group B residues displaying large values of local target function were moved to the group A and the minimization procedure was then reiterated. The distinction between an overall isotropic tumbling model (not requiring the use of molecule geometry) and an anisotropic model (using the atomic coordinates of the lowest energy NMR derived structure) was made based on Fisher–Snedecor statistics (F test).

3. Results

3.1. Sequence-specific assignments and secondary structure of the FCCH subdomain of mouse E1 enzyme

In the $^1H/^{15}N$ HSQC spectrum (Fig. 2) 104 out of 105 expected backbone correlations were identified. The exception was Lys304,

Table 1

NMR constraints and structural statistics for the ensembles of 20 lower energy (out from 200 submitted to calculations) of mouse FCCH conformers calculated and refined in explicit solvent in XPLOR-NIH 2.26 (Schwieters et al., 2003).

NOE distance constraints ^a	1558
Intra-residue ($ i - j = 0$)	330
Sequential ($ i - j = 1$)	457
Medium ($ i - j < 5$)	150
Long range ($ i - j > 5$)	621
Hydrogen bonds	38
Torsion angle constraints	
Backbone (ϕ/ψ)	146
Side chains (χ)	0
r.m.s.d. from experimental restraints	
Bonds (Å)	0.03988 ± 0.00168
Angles (deg)	0.50143 ± 0.10884
r.m.s.d. from idealized geometry	
Bond lengths (Å) · 10 ⁻³	2.422 ± 0.002
Bond angles (deg)	0.482 ± 0.001
Impropers (deg)	0.314 ± 0.001
Ramachandran plot (region 215..296) ^b	
Residues in most favored regions (%)	82.6
Residues in additional allowed regions (%)	15.1
Residues in generously allowed regions (%)	1.7
Residues in disallowed regions (%)	0.6
r.m.s.d. to the mean structure	
Ordered backbone atoms (217...293) (Å)	0.46 ± 0.09
Ordered heavy atoms (217...293) (Å)	1.01 ± 0.11
r.m.s.d. to the mean structure	
Ordered backbone atoms (201...312) (Å)	5.17 ± 1.29
Ordered heavy atoms (201...312) (Å)	5.75 ± 1.29
Structure Z-scores (217...293) ^c	
1st generation packing quality	1.616 ± 0.680
2nd generation packing quality	8.219 ± 2.141
Ramachandran plot appearance	-0.947 ± 0.484
χ^1/χ^2 rotamer normality	-2.919 ± 0.718
Backbone conformation	-1.499 ± 0.255

^a None of the 20 structures has a distance violation more than 0.5 Å and a dihedral angle violations more than 5°.

^b Quality of the NMR structures were checked by PROCHECK-NMR (v.3.4) program (Laskowski et al., 1996).

^c Ensemble of structures was validated and analyzed using WhatIF program (Vriend, 1990).

probably due to the correlation overlap of highly mobile residues clustering in the central region of the spectrum. The $^{13}C_\beta$ chemical shifts of three cysteines (Cys234, Cys262 and Cys278) showed that all cysteine thiol groups were in a reduced state (Sharma and Rajarathnam, 2000). Based on the $^{13}C_\beta$ and $^{13}C_\gamma$ chemical shifts, five out of six prolines (Pro216, Pro264, Pro272, Pro298 and Pro307) were in a *trans* conformation. Pro229 exhibited a *cis* conformation, confirmed by observation of cross peaks Pro229 H_α – Gly230 $H_{\alpha1,2}$ in the 3D NOESY-HSQC ^{13}C -edited spectrum. Secondary structure elements (six β -strands, one short 3_{10} -helix and one short α -helix) were initially deduced from H^N , N , C' , C_α and C_β chemical shifts predicted with PREDITOR server (Berjanskii et al., 2006) and then confirmed by observing characteristic cross peaks in 3D NOESY-HSQC spectra. The final positions of secondary structure motifs were estimated by STRIDE server (Heinig and Frishman, 2004) from the coordinates of the lowest energy structure after water refinement stage in XPLOR-NIH (Schwieters et al., 2003). The positions of six β -strands are as follows: Ser218–Thr225, Gly230–Cys234, Asp246–Gln254, Met265–Val269, Thr274–Ile277, Arg288–Gln294. A short α -helix and 3_{10} -helix comprise Ile257–Gly261 and Thr280–Asn282, respectively.

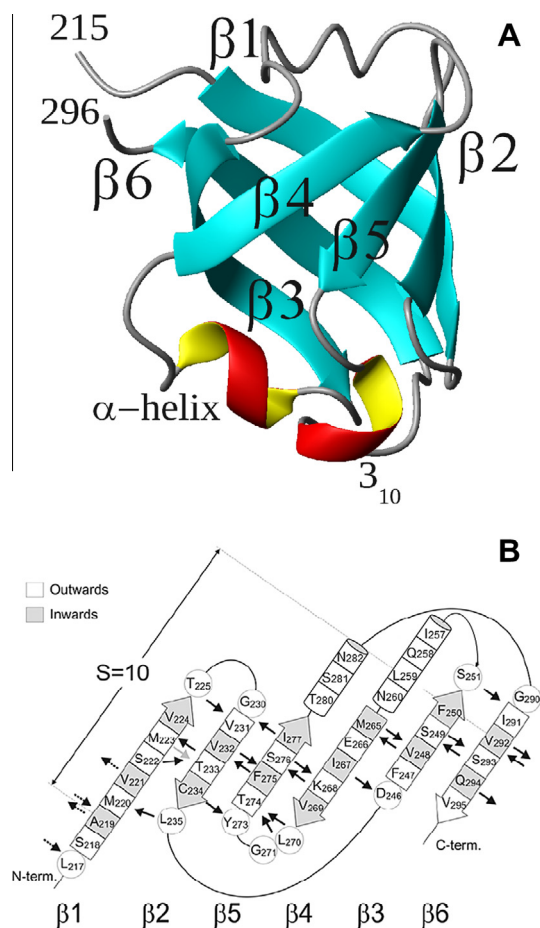


Fig. 3. (A) 3D high resolution solution structure of FCCH subdomain of mouse E1 enzyme. Its fold can be considered as a fusion of a β -hairpin (strands $\beta 1$ and $\beta 2$) and a Greek key motif (strands $\beta 5$, $\beta 4$, $\beta 3$ and $\beta 6$) that is interrupted by a very short, two turns: α -helix between strands $\beta 5$ and $\beta 6$ and by a 3_{10} -helix between strands $\beta 3$ and $\beta 4$. The figure was prepared in the MolMol program (Koradi et al. 1996) and secondary structure motifs were assigned by the STRIDE server (Heinig and Frishman, 2004). (B) Diagram of the secondary structure topology.

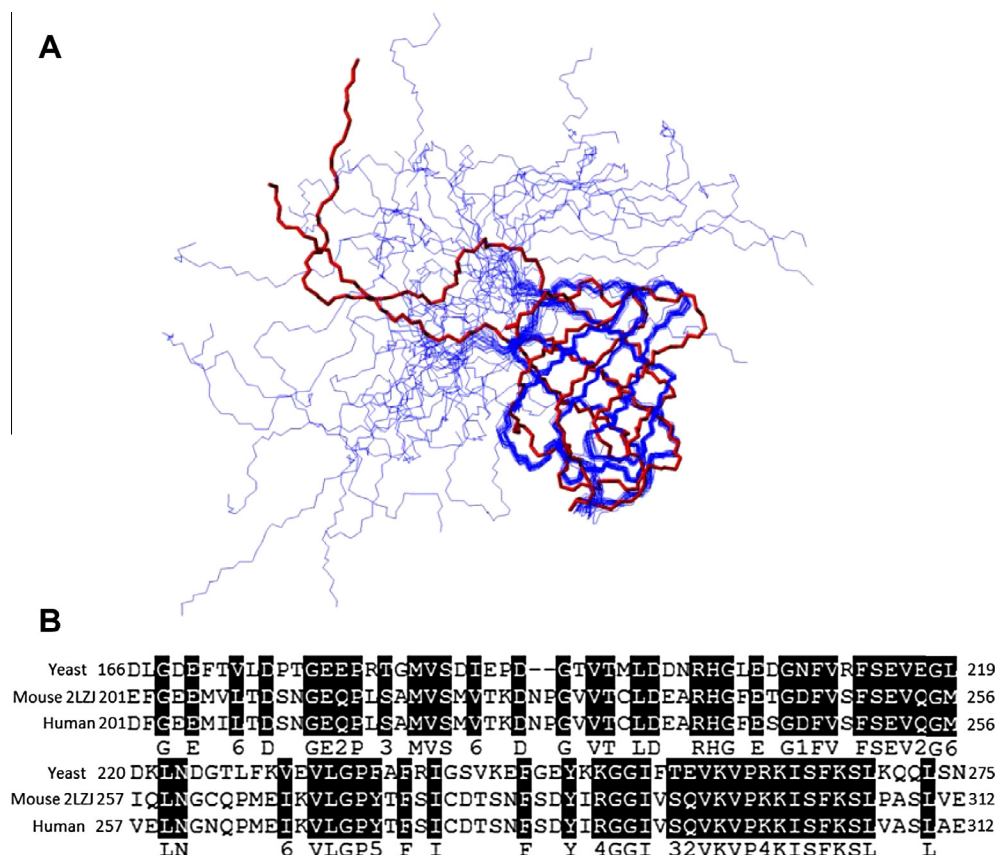


Fig. 4. (A) The superposition of the FCCH subdomains from yeast (red, X-ray structure) and mouse (blue, ensemble of 20 lowest energy NMR structures). The RMSD value calculated for the structured FCCH parts (NMR – residues 217–293; X-ray – residues 182–256) for the C α atoms is equal to 1.142 Å. (B) Alignment of FCCH domains of yeast, mouse and human Ubiquitin-E1s were obtained using MUSCLE (Edgar, 2004). Conserved residues are highlighted in black. The sequences share 58% identity. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.2. The FCCH subdomain of mouse E1 forms a six-stranded β -barrel

The three-dimensional structure of mouse FCCH has been determined based on NMR data (1558 distance constraints, 146 dihedral constraints derived from backbone chemical shifts, and 76 constraints resulting from hydrogen bonds). The statistics for the ensemble of the 20 most favorable FCCH structures is given in Table 1. Residues 202–216 and 294–312 are essentially unstructured in solution, at least if the FCCH fragment is isolated from the entire enzyme. In contrast, the core of the FCCH forms a well-defined six-stranded antiparallel β -barrel (Fig. 3). The details of the fold are the best understood from a schematic representation, which results from slicing the barrel between strands β 1 and β 6 and unwrapping it. β -strands are arranged in the order β 1, β 2, β 5, β 4, β 3, β 6 with +1, +3, –1, –1, +3 connectivity according to the Richardson nomenclature (Richardson, 1981). Formally, this fold can be considered as a fusion of a β -hairpin (strands β 1 and β 2) and a Greek key motif (strands β 5, β 4, β 3 and β 6), that is interrupted by a very short α -helix between strands β 3 and β 4 and one 3_{10} -helix between β 5 and β 6. One can conclude that the FCCH sequences and structures (Fig. 4) of mouse and yeast E1 enzymes (Lee and Schindelin, 2008) are similar.

3.3. ^{15}N relaxation-derived backbone dynamics of the FCCH subdomain

R_1 , R_2 and $\{^1\text{H}\}$ - ^{15}N NOE data have been obtained for 96 backbone amide groups out of 104 identified correlations. Data for the majority of residues located in both terminal parts contain negative NOEs and markedly smaller R_2 values characteristic for unstructured protein segments (Alexandrescu and Shortle, 1994; Zhukov

et al., 2007). In many cases, overlaps in 2D spectra of $^1\text{H}/^{15}\text{N}$ correlations preclude their quantitative analysis and therefore determination of relaxation parameters. Mostly the residues making up flexible termini are affected. The remaining data available for 69 residues of the central part of the construct (residues 217–293) differ from those for residues of both termini (Fig. 5). Their experimental relaxation parameters are rather uniform with only several outliers like $R_1 = 2.64 \pm 0.04 \text{ s}^{-1}$ (Gly255), $R_2 = 20.4 \pm 0.3 \text{ s}^{-1}$ (Glu237) or NOEs for Asp227 (0.28 ± 0.03), Val231 (0.41 ± 0.03), Asp236 (0.33 ± 0.03), and Ile267 (0.38 ± 0.02) in comparison with weighted mean values of relaxation data calculated for residues 217–293: 2.18 s^{-1} , 11.37 s^{-1} , and 0.66 for R_1 , R_2 and NOE, respectively. Initially, 44 residues were selected for the group exhibiting $R_{ex} = 0$ following the procedure described in Section 2.4. Finally, however, their number decreased to 27 (Fig. 6).

Comparison of two models of the overall tumbling (isotropic vs. fully anisotropic) favored the anisotropic model. The principal values of the anisotropic overall diffusion tensor are equal to: $D_x = (2.19 \pm 0.03)10^7 \text{ s}^{-1}$, $D_y = (1.36 \pm 0.02)10^7 \text{ s}^{-1}$, and $D_z = (1.86 \pm 0.03)10^7 \text{ s}^{-1}$. The averaged isotropic rotational correlation time, $\tau_R = (2D_x + 2D_y + 2D_z)^{-1} = 9.23 \pm 0.12 \text{ ns}$ is slightly larger than expected for a globular protein of 12.3 kDa size at 298 K. (Cavanagh et al., 2007), presumably due to the hindering effect of unstructured termini in our construct (Bae et al., 2009).

In the structured central part, the S^2 values are fairly uniform (mean S^2 value for residues located in the secondary structure elements is equal to 0.91 versus 0.89 for all residues). Only a few residues with diminished NOEs, situated in loops, display increased freedom of fast motions reflected by small S^2 values: Asp227 (0.68 ± 0.03), Val231 (0.72 ± 0.03) or Asp236 (0.80 ± 0.03) as shown

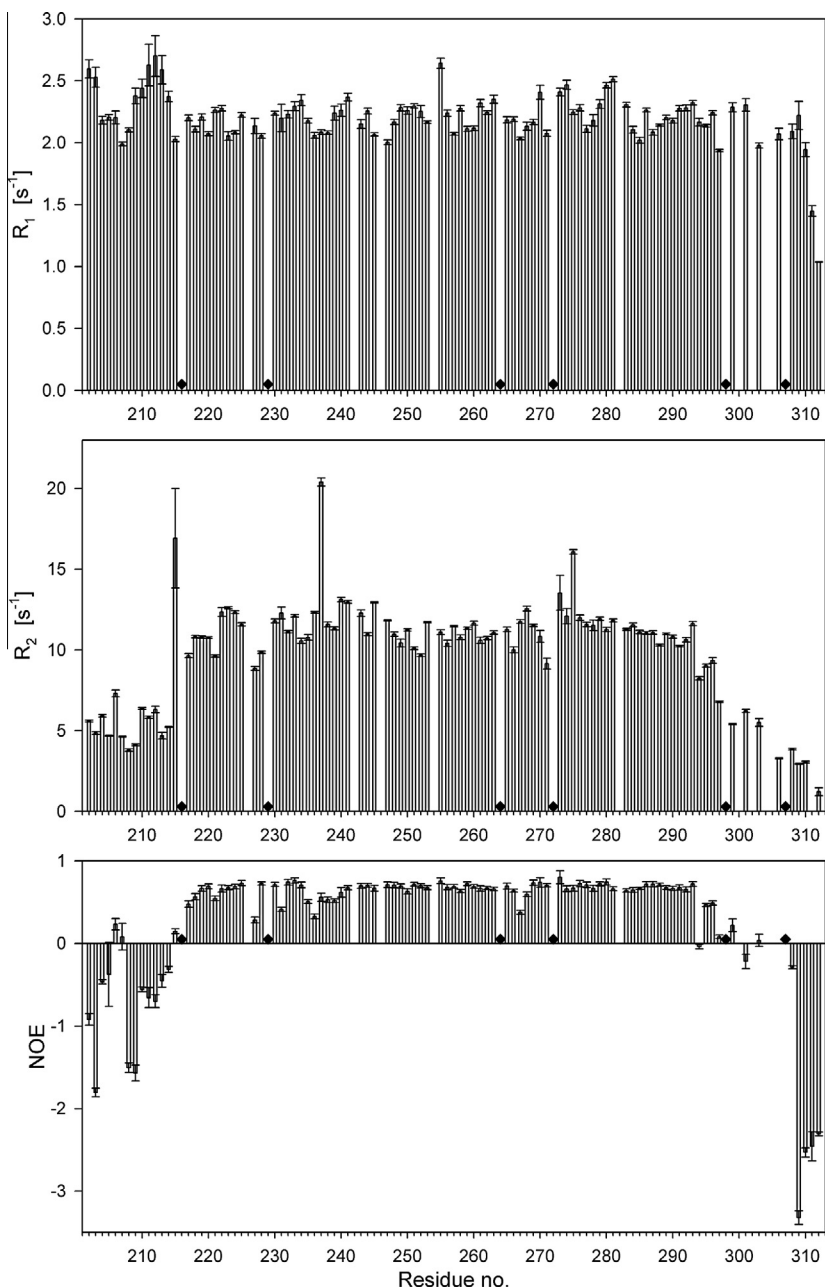


Fig. 5. Experimental values of ^{15}N backbone amide relaxation parameters R_1 , R_2 , and $(^1\text{H})\text{-}^{15}\text{N}$ NOE for the residues of the FCCH subdomain determined at 9.4 T. Asterisks mark the positions of Pro residues.

in Fig. 7. Long stretch of residues forming a loop between strands β_2 and β_3 (Leu235–Gly245) are characterized by intensive motions on the micro to millisecond time scale. Its mean R_{ex} value is equal to 2.65 s^{-1} in comparison to 1.13 s^{-1} for all residues and 0.86 s^{-1} for residues of secondary structure elements.

The obtained model parameters reproduce input experimental data very well ($r^2 > 0.99$). The plot of calculated vs. experimental R_1R_2 products is given in Supplementary Fig. S.1.

4. Discussion

4.1. Structure of the FCCH domain

The solution structure of the mouse FCCH subdomain presented in this work agrees closely (RMSD of structured part $\text{C}\alpha$ backbone atoms superposition equals 1.142 \AA) with the structured part of

the corresponding subdomain determined for the yeast E1 (Lee and Schindelin, 2008) (Fig. 4). A full turn around the 6-stranded β -barrel (in the direction of the hydrogen bonds) corresponds to a 10-residue shift around the barrel, typical for 6-stranded β -barrels (Murzin et al., 1994). The DALI scan (Holm and Rosenström, 2010) of the Protein Data Bank using FCCH as a search model identified β -barrel proteins with strand number $n = 6$ and shear number $S = 10$ as the most similar (Table 2). The selected proteins are involved in different cellular processes with no obvious connection to the Ub system. Therefore, the hits most likely show only structural similarity without functional implications.

4.2. Relaxation studies of the FCCH subdomain

The central fragment of the FCCH (217–293) is characterized by low backbone mobility typical for structured proteins, while unstructured C and N termini show much higher mobility (Figs. 5

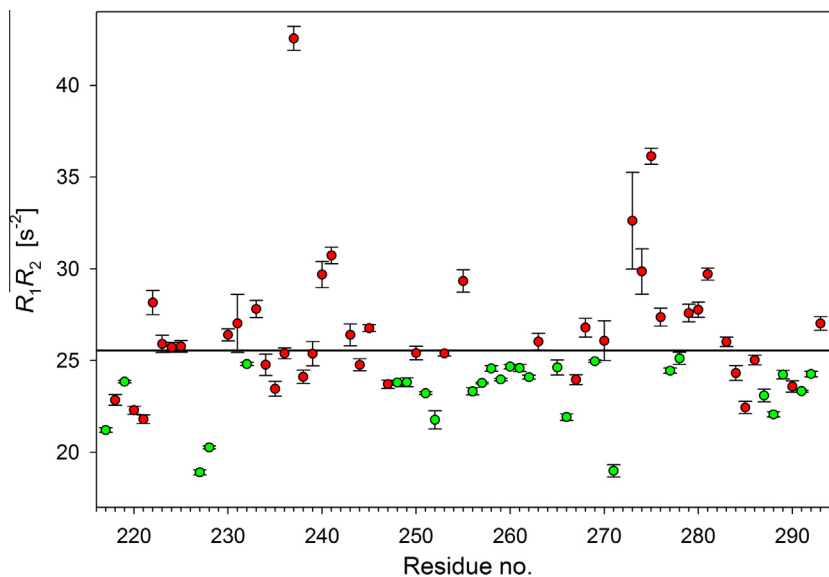


Fig. 6. R_1R_2 product obtained for the residues of the FCCH subdomain with corresponding error bars. The horizontal line represents the weighted mean value of R_1R_2 plus three standard deviations. All residues above this line were initially assumed to exhibit chemical exchange. The final partition between the residues affected by chemical exchange (red circles) and the residues which display $R_{ex} = 0$ (green circles) is color coded. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

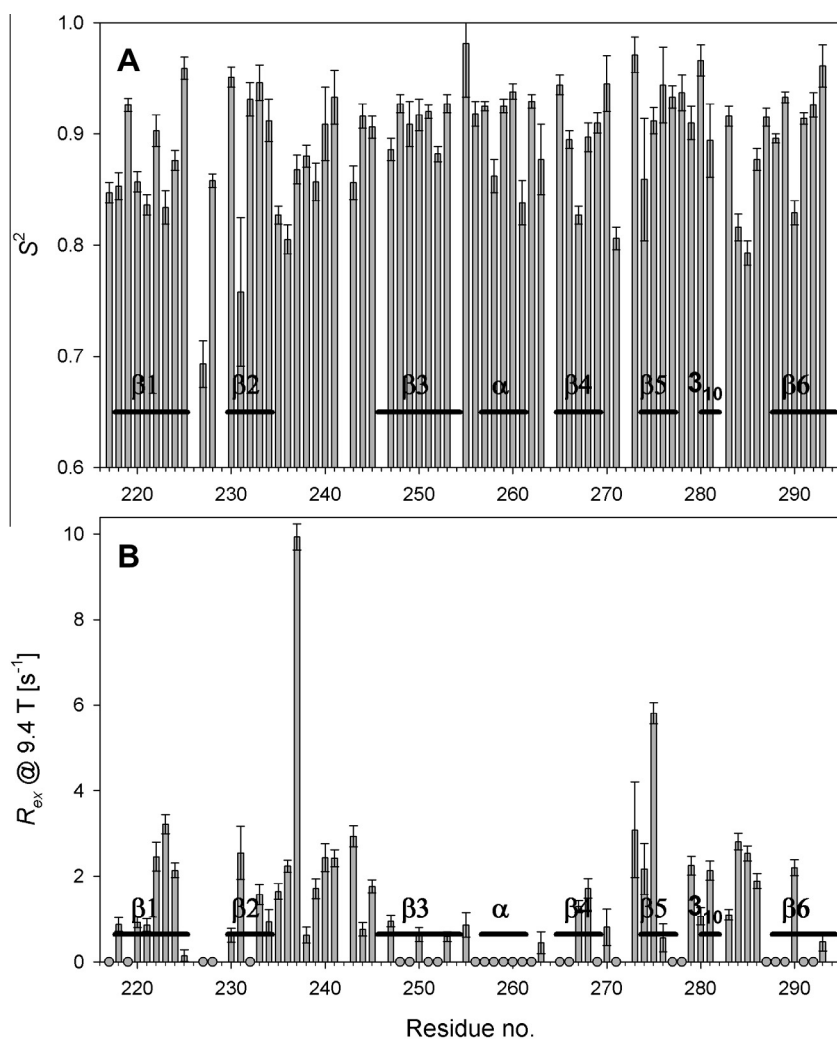


Fig. 7. Results of the model-free analysis of ^{15}N relaxation data for the structured part of the FCCH subdomain: data (vertical bars) for generalized order parameter values, S^2 (A) and chemical exchange, R_{ex} (B) with corresponding error bars. The horizontal lines indicate the positions of secondary structure elements. Residues for which no chemical exchange has been assumed are indicated with grey circles in the lower part of the figure.

Table 2

The DALI-score is a measure of structural similarity between two proteins in standard deviations above the statistically expected similarity.

Score	PDB	Function	Fold of the similar region
5.5	1QFJ	Flavin reductase	Barrel, closed; $n = 6, S = 10$; greek-key
5.4	1KK1	Initiation factor eIF2 gamma subunit	Barrel, closed; $n = 6, S = 10$; greek-key
5.4	1EFC	Elongation factor	Barrel, closed; $n = 6, S = 10$; greek-key – domain
5.3	1118	Riboflavin synthase	Barrel, closed; $n = 6, S = 10$; greek-key
4.9	2GPJ	Siderophore-interacting protein	Barrel, closed; $n = 6, S = 10$; greek-key
4.8	1AMO	NADPH-cytochrome p450 reductase	Barrel, closed; $n = 6, S = 10$; greek-key
4.7	2GLW	Transcription 92aa long hypothetical protein PHS018	Barrel, closed; $n = 6, S = 10$; greek-key
4.4	2GK7	Regulator of nonsense transcripts 1	Greek-key, $n = 6, S = 10(?)$
4.4	1SKY	F1-ATPase	Barrel, closed; $n = 6, S = 8$; greek-key
4.3	1G7R	Initiation factor IF2/eIF5b	Barrel, closed; $n = 6, S = 10$; greek-key
3.8	1FDR	Flavodoxin reductase ferredoxin reductase	Barrel, closed; $n = 6, S = 10$; greek-key
3.7	1TVC	Methane monooxygenase component c fragment	Barrel, closed; $n = 6, S = 10$; greek-key
3.6	1MRZ	Riboflavin kinase FMN adenyltransferase	Barrel, closed; $n = 6, S = 10$; greek-key
3.5	2PIA	Phthalate dioxygenase reductase	Barrel, closed; $n = 6, S = 10$; greek-key
3.5	2F1M	Acriflavine resistance protein A	Barrel, $n = 6, S = 8$ greek-key
3.5	1FNC	Ferredoxin reductase	Barrel, closed; $n = 6, S = 10$; greek-key
3.5	1SQR	50S ribosomal protein l35ae	Barrel, closed; $n = 6, S = 10$; greek-key
3.5	2FVG	Endoglucanase	Barrel, closed; $n = 6, S = 8$; greek-key
3.2	2D9R	Conserved hypothetical protein PG0164 from porphyromonas gingivalis [W83]	Barrel, closed; $n = 6, S = 10$; greek-key
3.2	1YOR	frv operon protein frvx	barrel, closed; $n = 6, S = 8$; greekkey

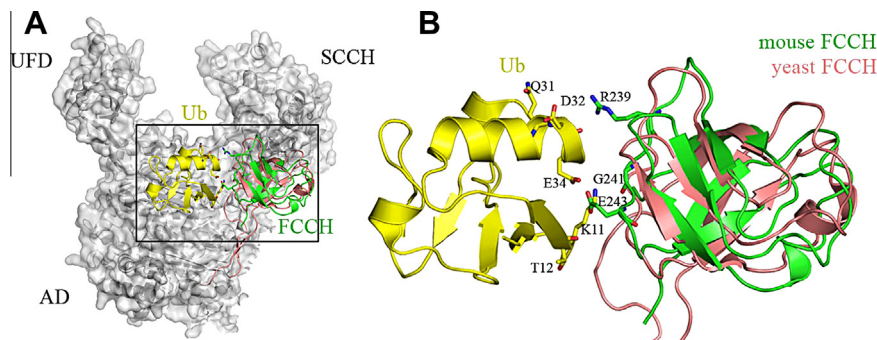


Fig. 8. Superimposition of the solution structure of FCCH of mouse E1 and the crystal structure of yeast E1-Ub complex. (A) Crystal structure of the yeast E1-Ub complex (PDB code: 3CMM). Yeast FCCH is shown in pink and ubiquitin in yellow. The FCCH from mouse E1 (green) has been superimposed on its yeast counterpart. (B) Close up view of interactions between FCCH and Ub. Side chains of residues which form hydrogen bonds between Ub and yeast E1 in the complex are shown in ball-and-stick representation.

and 7). Several residues located in loops are characterized by chemical exchange. Moreover, according to the crystal structure, the side chains of residues Arg202 (equivalent of Arg239 in mouse FCCH), Gly204 (equivalent of Gly241 in mouse FCCH) and Glu206 (equivalent of Glu243 in mouse FCCH) of yeast Ub-E1 form hydrogen bonds with Ub (cf. Fig. 8). All three residues are conserved and situated within the loop between $\beta 2$ and $\beta 3$ strands (Leu235–Gly245 region), in which residues exhibit elevated R_{ex} values, the hallmark of conformational exchange in the micro- to millisecond time scale. If the increased mobility could facilitate the adaptation of an appropriate interface conformation, the FCCH may have a role in the regulation of Ub binding to the E1 enzymatic machinery in the first stages of the Ub activation process, in agreement with an earlier suggestion that the FCCH may help to distinguish ubiquitin from other ubiquitin-like proteins (Viquez et al., 2012). This finding encouraged us to investigate the interaction of FCCH and Ub in solution. However, the FCCH titration with ubiquitin showed no visible chemical shift changes in the 2D $^1\text{H}/^{15}\text{N}$ HSQC spectra of FCCH. Hence, the FCCH in isolation (i.e. without the context of full length E1) does not bind to Ub in solution (Cavanagh et al., 2007).

4.3. Sequence alignment of the FCCH subdomain

FCCH domains of yeast, mouse and human Ub-E1s were aligned using MUSCLE (Edgar, 2004) (Fig. 4B). Conserved residues are highlighted in black. The sequences share 58% of identity. Such high

conservation in evolutionarily very distant species indicates significant selective pressure on this domain. Conserved residues are approximately equally distributed on the outside of the barrel, therefore no obvious protein–protein interfaces can be delineated from the structure of the FCCH fragment alone. In the yeast E1-Ub crystal structure FCCH forms one of the walls of the Ub adenylation pocket (Fig. 8). In the crystal, the side chains of Arg202 and Glu206 form hydrogen bonds with ubiquitin. In mouse E1 FCCH, both residues are situated within the Glu237 – E243 region, where residues exhibit conformational exchange in the micro to millisecond time scale. However, these residues are not among the most conserved residues in eukaryotes.

Inspection of the yeast E1 crystal structure (PDB id: 3CMM) (Lee and Schindelin, 2008) revealed that some contacts between FCCH and SCCH domains were present (Fig. S.2). FCCH and SCCH domains are evolutionary conserved and structurally similar. The backbone RMSD values between mouse catalytic half-domains and their yeast counterparts are 1.142 Å and 0.890 Å for FCCH and SCCH domain, respectively. These observations prompted us to test whether the mouse FCCH and SCCH domains interact in solution when not linked covalently by the adenylation domain. However, chemical shift perturbation in the $^1\text{H}/^{15}\text{N}$ HSQC spectrum of the FCCH/SCCH equimolar mixture was not observed. This finding points to a lack of interaction between these two catalytic half-domains in solution within the mM concentration range (Cavanagh et al., 2007). This result could either indicate that interactions are only observable when the two domains are covalently linked as

in the context of full-length E1, or could indicate that the interactions that are found in the crystal might be due to crystal packing.

5. Conclusions

The isolated FCCH domain adopts a native structure in aqueous solution. The structure of the FCCH subdomain of mouse E1 presented in this work closely resembles the structure of the corresponding subdomain of yeast E1 determined by Lee and Schindelin (2008). Interactions between SCCH and FCCH domains inferred from the crystal structure of E1 enzyme have not been confirmed by NMR spectroscopy. The lack of chemical shift perturbation in the $^1\text{H}/^{15}\text{N}$ HSQC spectrum of the FCCH/SCCH equimolar mixture might reflect the lack of covalent tethering in our experiments or a larger flexibility of full-length E1 than can be inferred from the crystal structure alone.

Accession numbers

The NMR resonance assignment for FCCH domain (residues 202–312) are available from the BMRB under accession number **18758**. Coordinates of the FCCH solution structure refined in explicit solvent in XPLOR-NIH are available from the Protein Data Bank (<http://www.rcsb.org>) with accession number 2LZJ.

Acknowledgments

We are grateful to H. Seino (National Institute of Genetics, Japan) for a clone of mouse E1 (Swiss-Prot Entry Q02053).

The authors would like to acknowledge financial support from: European Molecular Biology Organization (EMBO), Howard Hughes Medical Institute (HHMI) for a Young Investigator award, Foundation for Polish Science (START and VENTURES programs), EU European Social Fund, project number POKL.04.01.01-00-072/09, FP7 EAST-NMR, The Iuventus Plus project no. IP2011019371 from Polish Ministry of Sciences and Higher Education (L.J.), Polish Ministry of Higher Education grant N N301 318539 (I.Z.), Polish National Centre for Research and Development for a research grant number 178479 (contract number PBS1/A9/13/2012).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jsb.2013.10.020>.

References

- Alexandrescu, A.T., Shortle, D., 1994. Backbone dynamics of a highly disordered 131 residue fragment of staphylococcal nuclease. *J. Mol. Biol.* **242**, 527–546.
- Bae, S.H., Dyson, H.J., Wright, P.E., 2009. Prediction of the rotational tumbling time for proteins with disordered segments. *J. Am. Chem. Soc.* **131**, 6814–6821.
- Bax, A., Clore, G.M., Gronenborn, A.M., 1990. ^1H - ^1H correlation via isotropic mixing of ^{13}C magnetization: a new three dimensional approach for assigning ^1H and ^{13}C spectra of ^{13}C -enriched proteins. *J. Magn. Reson.* **88**, 425–431.
- Berjanskii, M.V., Neal, S., Wishart, D.S., 2006. *Nucleic Acids Res.* **34** (Web Server issue), W63–W69.
- Cavanagh, J., Fairbrother, W.J., Palmer III, A.G., Skelton, N.J., Rance, M., 2007. *Protein NMR Spectroscopy: Principles and Practice*, 2nd ed. Academic Press, Boston.
- Cooper, G.M., 2000. *The Cell: A Molecular Approach*, 2nd ed. Sinauer Associates, Sunderland (MA).
- Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., Bax, A., 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293.
- Edgar, R., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **5**, 1792–1797.
- EMBL Heidelberg. (n.d.). EMBL Protein Expression and Purification Core Facility. Retrieved from http://www.embl.de/pepcore/pepcore_services/protein_expression/ecoli/n15_c13/index.html.
- Fushman, D., 2003. Determination of protein dynamics using ^{15}N relaxation measurements. In: Zerbe, O. (Ed.), *BioNMR in Drug Research*. Wiley-VCH, Weinheim, Germany, pp. 283–308.
- Gardner, K.H., Konrat, R., Rosen, M.K., Kay, L.E., 1996. An (H)C(CO)NH-TOCSY pulse scheme for sequential assignment of protonated methyl groups in otherwise deuterated (^{15}N , ^{13}C -labeled) proteins. *J. Biomol. NMR* **8**, 351–356.
- Goddard, T.D., Kneller, D.G., 2010. SPARKY 3. University of California, San Francisco.
- Grzesiek, S., Bax, A., 1993. Amino acid type determination in the sequential assignment procedure of uniformly $^{13}\text{C}/^{15}\text{N}$ -enriched proteins. *J. Biomol. NMR* **3**, 185–204.
- Guentert, P., Mumenthaler, C., Wüthrich, K., 1997. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* **273**, 283–298.
- Haas, A.L., Warms, J.V., Hershko, A., Rose, I.A., 1982. Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation. *J. Biol. Chem.* **257**, 2543–2548.
- Heinig, M., Frishman, D., 2004. STRIDE: a Web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res.* **32**, W500–W502.
- Hershko, A., Ciechanover, A., 1998. The ubiquitin system. *Annu. Rev. Biochem.* **67**, 425–479.
- Holm, L., Rosenström, P., 2010. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–W549.
- Ikura, M., Kay, L.E., Bax, A., 1990. A novel approach for sequential assignment of ^1H , ^{13}C , and ^{15}N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochemistry* **29**, 4659–4667.
- Jaremko, L., Jaremko, M., Filipek, R., Wojciechowski, M., Szczepanowski, R.H., Bochtler, M., Zhukov, I., 2006. NMR assignment of a structurally uncharacterised fragment of recombinant mouse ubiquitin-activating enzyme. *J. Biomol. NMR* **36** (Suppl. 1), 43.
- Kay, L.E., Torchia, D.A., Bax, A., 1989. Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28**, 8972–8979.
- Kay, L.E., Keifer, P., Saareinen, T., 1992. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **114**, 10663–10665.
- Keller, R.L.J., 2004. Optimizing the process of nuclear magnetic resonance spectrum analysis and computer aided resonance assignment. Thèse de doctorat, ETH Zurich Thesis No. 15947, Switzerland.
- Kneller, J.M., Lu, M., Bracken, C., 2002. An effective method for the discrimination of rotational anisotropy and chemical exchange. *J. Am. Chem. Soc.* **124**, 1852–1853.
- Koradi, R., Billeter, M., Wüthrich, K., 1996. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics* **14**, 51–55.
- Laskowski, R.A., Rullmann, J.A.C., MacArthur, M.W., Kaptein, R., 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.
- Lee, I., Schindelin, H., 2008. Structural insights into E1-catalyzed ubiquitin activation and transfer to conjugating enzymes. *Cell* **134**, 268–278.
- Lipari, G., Szabo, A., 1982. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. *J. Am. Chem. Soc.* **104**, 4546–4570.
- Muhandiram, D.R., Farrow, N.A., Xu, G.-Y., Smallcombe, S.H., Kay, L.E., 1993. A gradient ^{13}C NOESY-HSQC experiment for recording NOESY spectra of ^{13}C -labeled proteins dissolved in H_2O . *J. Magn. Reson., Ser. B*, **317**–321.
- Muhandiram, D.R., Kay, L.E., 1994. Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. *J. Magn. Reson., Ser. B* **103**, 203–216.
- Murzin, A.G., Lesk, A.M., Chothia, C., 1994. Principles determining the structure of beta-sheet barrels in proteins. I A theoretical analysis. *J. Mol. Biol.* **236**, 1369–1381.
- Nowakowski, M., Jaremko, L., Jaremko, M., Zhukov, I., Belczyk, A., Bierzynski, A., Ejchart, A., 2011. Solution NMR structure and dynamics of human apo-S100A1 protein. *J. Struct. Biol.* **174**, 391–399.
- Nowakowski, M., Ruszczynska-Bartnik, K., Budzinska, M., Jaremko, L., Jaremko, M., Zdanowski, K., Bierzynski, A., Ejchart, A., 2013. Impact of calcium binding and thionylation of S100A1 protein on its nuclear magnetic resonance-derived structure and backbone dynamics. *Biochemistry* **52**, 1149–1159.
- Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T., 1986. *Numerical Recipes. The Art of Scientific Computing*. Cambridge University Press, Cambridge, Chapter 14.
- Richardson, J.S., 1981. The anatomy and taxonomy of protein structure. *Adv. Protein Chem.* **34**, 167–339.
- Schwieters, C.D., Kuszewski, J.J., Tjandra, N., Clore, G.M., 2003. The XPLOR-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65–73.
- Sharma, D., Rajarathnam, K., 2000. ^{13}C NMR chemical shifts can predict disulfide bond formation. *J. Biomol. NMR* **18**, 165–171.
- Stone, M.J., Fairbrother, W.J., Palmer 3rd, A.G., Reizer, J., Saier Jr., M.H., Wright, P.E., 1992. Backbone dynamics of the *Bacillus subtilis* glucose permease IIA domain determined from nitrogen-15 NMR relaxation measurements. *Biochemistry* **31**, 4394–4406.
- Szczepanowski, R.H., Filipek, R., Bochtler, M., 2005. Crystal structure of a fragment of mouse ubiquitin-activating enzyme. *J. Biol. Chem.* **280**, 22006–22011.
- Tjandra, N., Feller, S.E., Pastor, R.W., Bax, A., 1995. Rotational diffusion anisotropy of human ubiquitin from ^{15}N NMR relaxation. *J. Am. Chem. Soc.* **117**, 12562–12566.
- Viquez, O.M., Caito, S.W., McDonald, W.H., Friedman, D.B., Valentine, W.M., 2012. Electrophilic adduction of ubiquitin activating enzyme E1 by N, N-diethylthiocarbamate inhibits ubiquitin activation and is accompanied by striatal injury in the rat. *Chem. Res. Toxicol.* **25**, 2310–2321.

- Vriend, G., 1990. WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* 8, 52–56.
- Wittekind, M., Mueller, L., 1993. HNCACB, a high sensitivity 3D NMR experiment to correlate amide proton and nitrogen resonances with the α - and β -carbon resonances in proteins. *J. Magn. Reson. Ser. B* 101, 201–205.
- Woessner, D.E., 1962. Nuclear spin relaxation in ellipsoid undergoing rotational Brownian motion. *J. Chem. Phys.* 37, 647–654.
- Zhang, O., Kay, L.E., Olivier, J.P., Forman-Kay, J.D., 1994. Backbone ^1H and ^{15}N resonance assignments of the N terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J. Biomol. NMR* 4, 845–858.
- Zhukov, I., Bayer, P., Schoellermann, B., Ejchart, A., 2007. ^{15}N magnetic relaxation of backbone dynamics of the ribosome-associated cold shock response protein Yfia of *Escherichia coli*. *Acta Biochim. Pol.* 54, 769–775.