BIOLOGICAL CRYSTALLOGRAPHY

ISSN 1399-0047

Structural insights into the synthesis of FMN in prokaryotic organisms

Beatriz Herguedas,[a]‡ Isaias Lans,[a]§ María Sebastián,[a] Juan A. Hermoso,[b] Marta Martínez-Júlvez[a]* and Milagros Medina[a]*

[a]Departamento de Bioquímica y Biología Molecular y Celular, Facultad de Ciencias, and Instituto de Biocomputación y Física de Sistemas Complejos (Joint Unit BIFI–IQFR), Universidad de Zaragoza, Pedro Cerbuna 12, 50009 Zaragoza, Spain, and [b]GCMBE – Instituto Rocasolano, CSIC, Serrano 119, 28006 Madrid, Spain. *Correspondence e-mail: mmartine@unizar.es, medina@unizar.es

Received 17 July 2015
Accepted 16 October 2015

Edited by Z. S. Derewenda, University of Virginia, USA

‡ Present address: MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, England.
§ Present address: Grupo de Bioquímica Teórica, Universidad Industrial de Santander, Cra 27, Calle 9, Bucaramanga, Colombia.

Keywords: FAD synthetase; ATP:riboflavin kinase; substrate binding; conformational changes; crystal structure; molecular-dynamics simulations.

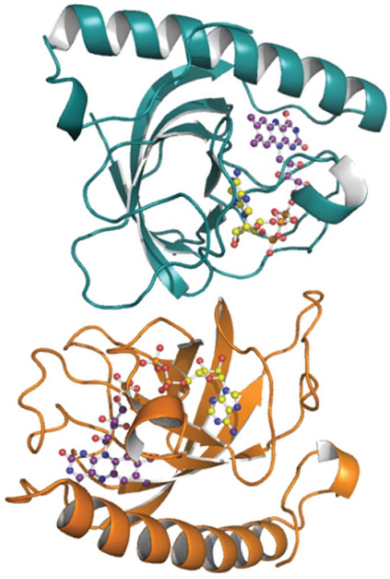
PDB references: RFK module of CaFADS, complex with ADP, 5a88; complex with FMN and ADP, 5a89; 5a8a

Supporting information: this article has supporting information at journals.iucr.org/d

Riboflavin kinases (RFKs) catalyse the phosphorylation of riboflavin to produce FMN. In most bacteria this activity is catalysed by the C-terminal module of a bifunctional enzyme, FAD synthetase (FADS), which also catalyses the transformation of FMN into FAD through its N-terminal FMN adenylyltransferase (FMNAT) module. The RFK module of FADS is a homologue of eukaryotic monofunctional RFKs, while the FMNAT module lacks homology to eukaryotic enzymes involved in FAD production. Previously, the crystal structure of *Corynebacterium ammoniagenes* FADS (*Ca*FADS) was determined in its apo form. This structure predicted a dimer-of-trimers organization with the catalytic sites of two modules of neighbouring protomers approaching each other, leading to a hypothesis about the possibility of FMN channelling in the oligomeric protein. Here, two crystal structures of the individually expressed RFK module of *Ca*FADS in complex with the products of the reaction, FMN and ADP, are presented. Structures are complemented with computational simulations, binding studies and kinetic characterization. Binding of ligands triggers dramatic structural changes in the RFK module, which affect large portions of the protein. Substrate inhibition and molecular-dynamics simulations allowed the conformational changes that take place along the RFK catalytic cycle to be established. The influence of these conformational changes in the FMNAT module is also discussed in the context of the full-length *Ca*FADS protomer and the quaternary organization.

1. Introduction

The essential cofactors of flavoproteins and flavoenzymes, flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD), are synthesized from riboflavin (RF; vitamin B_2) in two sequential reactions: RF phosphorylation is catalysed by an ATP-riboflavin kinase (RFK; EC 2.7.1.26) to produce FMN, which can be then converted to FAD by an FMN:ATP adenylyltranferase (FMNAT; EC 2.7.7.2) (Efimov *et al.*, 1998; Barile *et al.*, 2000; Giancaspero *et al.*, 2009; Sandoval *et al.*, 2008). The enzymes catalysing these activities are very diverse, with different protein folds encoding the same activities, as well as monofunctional and bifunctional members. In general mammals and yeast have two independent monofunctional enzymes that catalyse each of the reactions, while bacteria contain a single bifunctional polypeptide called FAD synthetase (FADS) (Serrano, Ferreira *et al.*, 2013; Frago *et al.*, 2008; Yruela *et al.*, 2010; Krupa *et al.*, 2003). Whereas the N-terminal module of FADS lacks structural homology to eukaryotic FMNATs, the kinase module is homologous to monofunctional RFKs (Herguedas *et al.*, 2010; Huerta *et al.*, 2009). Archaea have a CTP-dependent RFK, as well as an air-sensitive



© 2015 International Union of Crystallography

enzyme that transforms FMN into FAD and is not related to either eukaryotic or prokaryotic FMNATs (Mashhadi *et al.*, 2008, 2010). Finally, plants have a complex distribution of these activities, with prokaryotic-like and eukaryotic-like modules associated with others or apparently degenerated (Yruela *et al.*, 2010; Sandoval & Roje, 2005; Sandoval *et al.*, 2008).

The structural information on these enzymes is increasing slowly (Supplementary Table S1; Serrano, Ferreira *et al.*, 2013; Medina, 2012; Herguedas *et al.*, 2010; Leulliot *et al.*, 2010; Wang *et al.*, 2005; Karthikeyan, Zhou, Mseeh *et al.*, 2003; Karthikeyan, Zhou, Osterman *et al.*, 2003), but the structures of plant proteins, human FADS and the air-sensitive archaeal FADS remain elusive. We have previously solved the crystal structure of the *Corynebacterium ammoniagenes* FAD synthetase (*Ca*FADS; Herguedas *et al.*, 2010). Like other prokaryotic FADSs, *Ca*FADS folds into two almost independent modules: an N-terminal module (residues 1–186) catalysing the FMNAT reaction and a C-terminal module (187–338) that contains the RFK activity (Frago *et al.*, 2008, 2009; Karthikeyan, Zhou, Osterman *et al.*, 2003; Bauer *et al.*, 2003; Herguedas *et al.*, 2010; Supplementary Fig. S1*a*). Based on the characterization of different protein mutants, it was established that residues Thr208 and Asn210 (*Ca*FADS numbering) in the consensus 207-PTAN-210 motif of RFK provide the active-site geometry for substrate binding and catalysis, while Glu268, which is also conserved, acts as a catalytic base during FMN biosynthesis (Supplementary Fig. S2*a*; Serrano, Frago *et al.*, 2013; Bauer *et al.*, 2003; Karthikeyan, Zhou, Mseeh *et al.*, 2003; Karthikeyan, Zhou, Osterman *et al.*, 2003; Cheek *et al.*, 2005; Frago *et al.*, 2008). The RFK activity of *Ca*FADS exhibits inhibition by excess RF, as also described for the monofunctional rat liver RFK, which is apparently related to the creation of nonproductive enzyme–substrate/product complexes, the nature of which is still not clear (Yamada *et al.*, 1990; Serrano, Frago *et al.*, 2013). Overall, the *Ca*FADS RFK module has high structural homology to eukaryotic RFKs (such as those from *Schizosaccharomyces pombe*, *Trypanosoma brucei* and *Homo sapiens*; *Sp*RFK, *Tb*RFK and *Hs*RFK, respectively; Supplementary Fig. S2*a*); however, while the ATP-binding site in the apo RFK module of *Ca*FADS is closed by several loops, the apo forms of these monofunctional enzymes present an open binding site (Herguedas *et al.*, 2010; Bauer *et al.*, 2003; Karthikeyan, Zhou, Mseeh *et al.*, 2003; Karthikeyan, Zhou, Osterman *et al.*, 2003; Supplementary Figs. S2*b* and S2*c*). Additionally, the conformation of Pro207 and Thr208 in the PTAN motif differs from that in monofunctional RFKs, with Thr208 blocking the binding site of the adenine moiety of ATP in *Ca*FADS (Karthikeyan, Zhou, Mseeh *et al.*, 2003; Bauer *et al.*, 2003; Serrano, Frago *et al.*, 2013; Supplementary Fig. S2*d*).

The crystal structure of *Ca*FADS is consistent with the formation of a dimer of trimers in which neighbouring protomers interact by the approach of their RFK and FMNAT active sites (Herguedas *et al.*, 2010; Serrano, Frago *et al.*, 2013). The formation of transient hexameric oligomers was also observed in solution by atomic force microscopy (AFM) and detected at a cellular level. Despite the fact that the monomer appears to be the major form in solution, these studies showed a correlation between ligand binding at the RFK module and stabilization of the quaternary organization; more recently, some of the residues responsible for this organization have also been identified (Marcuello *et al.*, 2013; Serrano *et al.*, 2015; Supplementary Fig. S1*b*).

In this context, it is of relevance to determine the molecular mechanism of FMN biosynthesis in *Ca*FADS. We focused on the crystallization of protein–ligand complexes using the isolated RFK module of *Ca*FADS, Δ(1–182)*Ca*FADS. Here, we present its crystal structure in complex with ADP-Ca^{2+} (referred to here as the binary complex) and with FMN–ADP-Mg^{2+} (referred to as the ternary complex) together with molecular-dynamics simulations and enzymatic characterization. Analysis of these structures provides relevant insights into the structure–function dependence of the RFK activity in this prokaryotic FADS.

2. Materials and methods

2.1. Cloning, expression and purification

The truncated gene coding for residues 183–338 of *Ca*FADS, Δ(1–182)*Ca*FADS, and corresponding to the RFK module, was cloned into the pET-28a plasmid, transformed into *Escherichia coli* BL21(DE3) cells and expressed and purified as described previously (Frago *et al.*, 2009). The pure protein was concentrated to 10 mg ml^{-1} in 20 m*M* Tris–HCl pH 8.0 with 150 m*M* NaCl.

2.2. Crystallization

Initial crystallization screenings were performed by the sitting-drop vapour-diffusion method in 96-well plates containing commercial screens from Hampton Research and Jena Biosciences at 18°C. Prior to crystallization, a solution of the protein at 7.5–10 mg ml^{-1} was incubated with 1 m*M* $MgCl_2$, 1 m*M* FMN and/or 1 m*M* ADP for 20 min at 4°C. Trials with other protein ligands such as RF or ATP have also been carried out, but crystals have not been obtained so far.

The ternary complex Δ(1–182)*Ca*FADS–FMN–ADP-Mg^{2+} crystallized in several conditions containing PEG 6000, PEG 4000 or PEG 3350 in 100 m*M* sodium acetate pH 4.6–5.0 and different salts. Crystals were further optimized using Additive Screen from Hampton Research. Diffraction-quality crystals appeared in drops consisting of 2 µl 26–30% PEG 4000, 200 m*M* Li_2SO_4, 100 m*M* sodium acetate pH 5.0, 2 µl protein solution and 0.2 µl 1 *M* NaI equilibrated against 500 µl mother liquor. Two crystal forms were obtained: crystal TI, with dimensions of 0.15 × 0.05 × 0.05 mm, and crystal TII, with dimensions of 0.15 × 0.15 × 0.15 mm.

The binary complex Δ(1–182)*Ca*FADS–ADP-Mg^{2+} crystallized using different PEGs (PEG 3350, PEG MME 5500, PEG 6000 and PEG 8000) in 0.1 *M* MES pH 6.5 and a salt of a divalent cation (magnesium formate, magnesium acetate, magnesium chloride or calcium chloride). Diffraction-quality

Table 1
Data-collection statistics, processing and refinement statistics.

Values in parentheses are for the highest resolution shell.

	Δ(1–182)*Ca*FADS–FMN–ADP-Mg^{2+}		Δ(1–182)*Ca*FADS–ADP-Ca^{2+}	
	Crystal TI	Crystal TII	Crystal BI	Crystal BII
Data collection and processing				
Wavelength (Å)	0.93300	0.93300	0.97600	0.97600
Temperature (K)	100	100	100	100
Space group	$P2_12_12_1$	$P3_221$	$P3_2$	$P3_2$
Unit-cell parameters				
a (Å)	43.40	73.59	68.66	68.75
b (Å)	70.48	73.59	68.66	68.75
c (Å)	100.08	149.69	147.00	148.20
$\alpha = \beta$ (°)	90	90	90	90
γ (°)	90	120	120	120
Resolution range (Å)	35.24–1.65 (1.74–1.65)	36.79–1.80 (1.90–1.80)	46.22–2.08 (2.19–2.08)	59.54–2.15 (2.27–2.15)
Total No. of reflections	190705	193808	260610	144755
No. of unique reflections	36054	43268	46431	40688
$\langle I/\sigma(I)\rangle$	18.5 (5.1)	15.8 (4.0)	11.2 (3.8)	7.1 (2.4)
Completeness (%)	95.7 (76.6)	98.0 (89.6)	99.7 (100)	95.6 (99.7)
Multiplicity	5.3 (3.0)	4.5 (3.8)	5.6 (5.5)	3.6 (3.5)
R_{merge}	0.05 (0.16)	0.06 (0.20)	0.10 (0.44)	0.10 (0.46)
CC_Imean†	99.8 (94.9)	99.7 (92.9)	99.2 (86.2)	
Twin fraction			0.214	0.5
Refinement				
R_{work} (%)	14.6	17.3	15.9	—
R_{free} (%)	17.5	20.3	18.9	—
No. of protomers in asymmetric unit	2	2	4	—
No. of non-H atoms				
Protein	2368	2420	4361	—
$Mg^{2+}/Zn^{2+}/Ca^{2+}$	4	4	4	—
FMN	62	62	—	—
ADP	54	54	108	—
Solvent	382	324	295	—
R.m.s.d., bonds (Å)	0.015	0.012	0.012	—
R.m.s.d., angles (°)	1.772	1.477	1.584	—
Average B factor (Å^2)				
Protein	19/14‡	20/19‡	48/50/52/59§	—
FMN	16/10‡	15/12‡	—	—
ADP	12/8‡	15/13‡	32/35/37/47§	—
Mg^{2+}	6/3‡	11/8‡	—	—
Zn^{2+}	13/19‡	17/21‡	—	—
Ca^{2+}	—	—	33/33/35/45§	—
PDB code	5a89	5a8a	5a88	—

† Equivalent to $CC_{1/2}$ (labelled 'CC_Imean' in older versions of *SCALA*; Evans & Murshudov, 2013). ‡ Chain *A*/chain *B*. § Chain *A*/chain *B*/chain *C*/chain *D*.

crystals grew in 10–14% PEG 8000, 20% glycerol, 0.1 *M* MES–NaOH pH 6.5, 200 m*M* $CaCl_2$.

2.3. Data collection, structure solution and refinement

Diffraction data for the ternary complex were collected on the ID14.1 beamline (λ = 0.9330 Å) at the European Synchrotron Radiation Facility (ESRF), Grenoble, obtaining two data sets to 1.65 Å (crystal TI) and 1.80 Å (crystal TII) resolution. Data for the binary complex were collected on the ID23.1 beamline (λ = 0.97600 Å) at the ESRF, yielding two data sets to maximum resolutions of 2.08 Å (crystal BI) and 2.15 Å (crystal BII). The data sets were processed with *iMosflm* (Battye *et al.*, 2011) and were scaled and reduced with *SCALA* (Evans, 2006) from the *CCP*4 package (Table 1).

Crystal TI belonged to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters $a = 43.4$, $b = 70.5$, $c = 100.1$ Å, and crystal TII belonged to the trigonal space group $P3_221$, with unit-cell parameters $a = b = 73.6$, $c = 149.7$ Å. The binary complex crystals BI and BII had similar unit-cell parameters to crystal TII ($a = b = 68.7$, $c = 147.0$ Å for BI and $a = b = 68.7$, $c = 148$ Å for BII), but belonged to space group $P3_2$ (Table 1).

2.3.1. Structure determination of the ternary complex. The structure of the ternary complex was solved by molecular replacement (MR) with *Phaser* (Vagin & Teplyakov, 2010; Navaza, 2001; McCoy *et al.*, 2007) using the RFK module of *Ca*FADS as a search model (PDB entry 2x0k; Herguedas *et al.*, 2010) and the diffraction data of crystal TI. Interestingly, we could only obtain a MR solution after removing 46 residues of the model (L1c-FlapI, L3c and L4c-FlapII; Supplementary Fig. S2) and only with crystal TI. The initial maps clearly showed density for FMN and ADP-Mg^{2+} in the active site as well as density for the removed residues. The model was manually rebuilt using *Coot* (Emsley & Cowtan, 2004) and refinement was performed with *REFMAC*5 (Murshudov *et al.*, 2011) using noncrystallographic symmetry restraints and TLS refinement. The final model includes residues 186–336 in chain *A* and 186–338 in chain *B*, and one molecule of FMN, ADP-Mg^{2+} and Zn^{2+} ion in each chain, as well as 382 water molecules.

The structure of crystal TII was solved using the partially refined coordinates of crystal TI as an initial model. Although the unit-cell parameters suggested the presence of three molecules in the asymmetric unit (V_M of 2.28 Å^3 Da^{-1} and solvent content of 46%, similar to those of crystal TI), MR indicated that only two protein chains compose the asymmetric unit (V_M of 3.42 Å^3 Da^{-1} and solvent content of 64%). The refined model contains residues 184–337 in chain *A* and residues 185–338 in chain *B*, 324 waters, two molecules of FMN, ADP-Mg^{2+} and the additional Zn^{2+} ions observed in the crystal TI.

Both models were validated with *MolProbity* (Chen *et al.*, 2010).

2.3.2. Structure determination of the binary complex: twinning. The structure of the binary complex was solved with

MOLREP (Vagin & Teplyakov, 2010) using the coordinates of the ternary complex as a search model. MR fixed the ambiguity in the space-group determination, indicating that the correct space group was $P3_2$ with four molecules in the asymmetric unit. The high R and R_{free} values observed during refinement suggested that the crystals were twinned. Analysis with the Padilla–Yeates server (http://nihserver.mbi.ucla.edu/pystats/; Padilla & Yeates, 2003) indicated partial merohedral twinning for crystal BI and perfect twinning for crystal BII. Since the structure had been solved in the correct space group regardless of the twinning, we carried out the refinement using the original diffraction data and the twin-refinement option of *REFMAC*5, refining against intensities rather than amplitudes. As a result, the R and R_{free} values decreased to 18 and 20%, respectively.

The final model consists of four peptide chains, four ADP molecules, four Ca^{2+} ions, six glycerol molecules and 295 waters. Ca^{2+} ions are present in the crystallization condition at 0.2 *M*, displacing Mg^{2+} in the coordination of ADP. Several protein loops had poor electron density and were not included in the final model. Thus, residues 258–263 (L4c-FlapII) are not visible in any of the chains, while residues 325–330 (L7c) show poor density in chains *A* and *D* and are not visible in chains *B* and *C*. Finally, residues 197–203 (L1c-FlapI) are only observed in chains *C* and *D*. Table 1 lists the refinement statistics for crystal BI.

2.4. Enzymatic characterization

The RFK activity was measured at 25°C in a final volume of 500 µl in 20 m*M* PIPES, 0.8 m*M* $MgCl_2$ pH 7.0 containing 0.3–18 µ*M* RF, 4–300 µ*M* ATP as described previously (Serrano, Frago *et al.*, 2013; Serrano *et al.*, 2012). The reaction was initiated by the addition of ∼35 n*M* of the monomeric enzyme followed by 1 min of incubation at 25°C and was stopped at 100°C for 5 min. The composition of flavins in the supernatant was analysed using an Alliance HPLC system (Waters) equipped with a 2707 autosampler and a HSST3 column (4.6 × 150 mm, 3.5 µm; Waters) preceded by a pre-column (4.6 × 20 mm, 3.5 µm; Waters) as described previously (Serrano, Frago *et al.*, 2013; Serrano *et al.*, 2012). Experiments were performed in triplicate. The kinetic data obtained at different RF concentrations with ATP at saturating concentrations (as nanomoles of flavin transformed per minute) showed inhibition by excess RF; therefore, a model describing the substrate-inhibition effect produced in bi-substrate enzyme kinetics was used to interpret the experimental data and to obtain kinetic parameters (Leskovac, 2003). The kinetic data obtained for ATP at RF concentrations producing the maximal detected activity were interpreted using the Michaelis–Menten kinetic model.

2.5. High-sensitivity isothermal titration calorimetry (ITC)

Measurements were carried out using an AutoITC200 calorimeter (MicroCal) thermostated at 25°C. Typically, 180 µ*M* RF and 300 µ*M* ATP solutions were used to titrate ∼25 µ*M* of the apo RFK module in a 200 µl cell volume. Titrations of the RFK module saturated with either ATP (RFK–ATP) or RF (RFK–RF) were carried out by stepwise injections of the second ligand solution (RF or ATP, respectively) into a calorimetric cell containing a mixture of the protein with either a saturating concentration of ATP (400–500 µ*M*) or 180 µ*M* RF. Ligands and protein were dissolved in 20 m*M* PIPES pH 7.0 and degassed prior to titration. All determinations presented here were carried out in the absence of Mg^{2+} to avoid catalysis. Up to 19 injections of 2 µl were added to the cell sample and mixed using a rotating (1000 rev min^{-1}) stirrer syringe. The association constant (K_a), the enthalpy change (ΔH) and the stoichiometry (N) were obtained through nonlinear regression of the experimental data to a home-derived model for one or two independent binding sites implemented in *Origin* 7.0 (OriginLab) as described previously (Serrano, Frago *et al.*, 2013; Serrano *et al.*, 2012). The dissociation constant (K_d), the free-energy change (ΔG) and the entropy change (ΔS) were obtained from basic thermodynamic relationships. Experiments were performed in duplicate or triplicate. The errors considered in the measured parameters (±15% in K_d and ±0.3 kcal mol^{-1} in ΔH and $-T\Delta S$) were taken to be larger than the standard deviation between replicates and the numerical error after fitting analysis.

2.6. Molecular-dynamics (MD) simulations

We used the Cartesian coordinates of residues 187–338 of the crystal structure of free *Ca*FADS (PDB entry 2x0k) after the removal of crystallographic waters (referred to here as free RFK) as the starting structure for the MD simulations. Models of different complexes were obtained by superposition of the ternary complex on this free RFK structure or on those derived by MD simulations.

MD simulations were carried out using *AMBER* (Case *et al.*, 2005), ff94 (Cornell *et al.*, 1995) and the parameters for FMN, RF, ADP and ATP (http://www.pharmacy.manchester.ac.uk/bryce/amber#cof; Antony *et al.*, 2000; Walker *et al.*, 2002). Residues were protonated at pH 7.0 using standard *AMBER* methods and *PROPKA* (Bas *et al.*, 2008). Each molecular system was neutralized by the addition of sodium ions and solvated with a TIP3P water model in a truncated octahedral box centred at the geometric centre of the protein and with a minimum distance of 10 Å to any protein atom. To remove close contacts and highly repulsive orientations of the initial protein–solvent system, solvent molecules and counter-ions were relaxed and allowed to redistribute around the restrained protein molecule before energy minimization of the entire system. The process consisted of 5000 steps of steepest descent (SD) and 5000 steps of conjugate gradient (CG), in which the protein was restrained with an elastic force of 500 kcal mol^{-1} $Å^{-2}$ centred at each atom to smooth the solvent–protein interaction. A second energy minimization applied 5000 steps of SD and 5000 steps of CG to the system without restraints. We heated the resulting system from 0 to 300 K with the protein atoms weakly constrained with a harmonic force constant of 10 kcal mol^{-1} $Å^{-2}$ for 40 ps to

obtain a solvent pre-equilibration. The whole system was then equilibrated for 300 ps at 300 K and at a constant pressure of 101.3 kPa with a weak-coupling pressure algorithm. We used a Langevin temperature-equilibration algorithm, periodic boundary conditions and the particle mesh Ewald method for long-range electrostatic interactions. All bonds involving H atoms were constrained with *SHAKE*. A time step of 2 fs was used for dynamics integration. Each system was simulated for 1–15 ns, maintaining the same conditions as for the second equilibration, and structures were collected every 2 ps. During equilibration and simulations we used the leapfrog Verlet integration scheme.

2.6.1. Targeted molecular-dynamics simulations. Targeted MD is a useful method to induce conformational changes to reach a known target structure by the application of an additional term to the energy function (Schlitter *et al.*, 1994). This extra term is based on the mass-weighted r.m.s.d. of a set of atoms in the structure under study compared with the target structure that we want to reach. This energy term has the form $E = 0.5 \times \mathrm{TF} \times \mathrm{TFR} \times [\mathrm{r.m.s.d.}(t) - \mathrm{TGR}(t)]^2$, where TF is the force constant for targeted MD, TFR defines the atoms that will be used for the r.m.s. superposition between the current structure and the target structure, the time-dependent r.m.s.d.(t) is the r.m.s.d. between the selected atoms of the current structure and the target structure, and the time-dependent TGR(t) is the value of the target r.m.s.d. The force constant used was 1 kcal mol^{-1} Å^{-2} and the length of each target dynamics simulation was 1 ns.

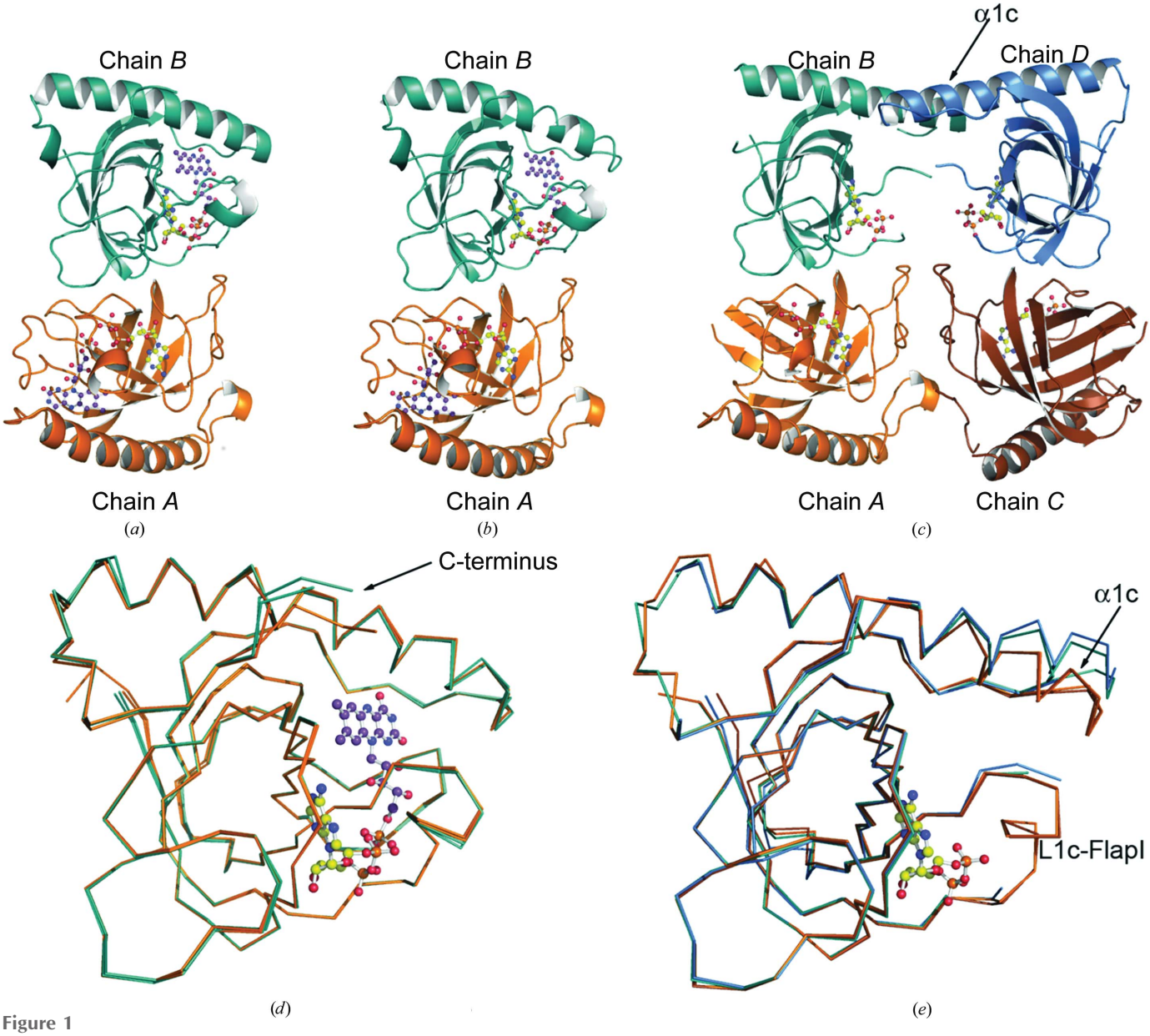


Figure 1
Crystal structures of the RFK module of binary and ternary complexes of *Ca*FADS. Cartoon representations of the asymmetric unit of the ternary complex in (*a*) crystal TI and (*b*) crystal TII and (*c*) of the binary complex. (*d*) Superposition of chains *A* and *B* of crystal TI, showing the main differences in the C-terminus of the protein sequence. (*e*) Superposition of the four chains of the binary complex, showing the main differences at the α1c helix, which is distorted in chains *B* and *D* owing to crystal contacts.

We analysed MD trajectories by inspection of the evolution of the three-dimensional conformation using *VMD* (Humphrey *et al.*, 1996) and *PyMOL* (DeLano, 2002). Interatomic distances and angles, as well as the r.m.s.d. from a given structure, were monitored using the *ptraj* module in *AMBER* (Case *et al.*, 2005). The secondary-structure information along the MD was calculated following the *Database of Secondary Structure in Proteins* (*DSSP*) method (Kabsch & Sander, 1983*a*,*b*).

3. Results

3.1. Overall structures of the ternary and binary complexes

Crystals TI and TII of the ternary complex contain two protein chains in the asymmetric unit (Figs. 1*a* and 1*b*). The four monomers are very similar (r.m.s.d. on C^{α} atoms of below 0.25 Å), showing subtle differences in the C-terminal loop (Fig. 1*d*). Here, we will use chain *A* of crystal TI to describe the structure of the ternary complex. The crystal of the binary complex has four molecules in the asymmetric unit, grouped into two pairs (*AB* and *CD*) with an arrangement similar to that observed in the ternary complex (Fig. 1*c*). Chains *A* and *C* differ from chains *B* and *D* in the conformation of the C-terminal helix (α1c) and in residues 197–203 (L1c-FlapI), which are absent in the *A* and *B* protomers (Fig. 1*e*). The α1c helix is distorted in chains *B* and *D* owing to crystal contacts that do not occur in chains *A* and *C* (Fig. 1*e*). Thus, chains *A* and *C* will be used to describe the structure of the binary complex. Although both the binary and the ternary complexes show a similar dimeric arrangement, analysis of the crystal contacts with *PISA* suggests that the dimers are not stable in solution.

In both structures the protein adopts the typical fold of the RFK family, consisting of a central β-barrel with six anti-parallel β-strands, a long α-helix located at the C-terminus (α1c) and several loops connecting the β-strands (L1c-FlapI, L2c, L3c, L4c-FlapII, L5c, L6c and L7c) (Fig. 2; Supplementary Fig. S2). Although crystal contacts are observed in these loops among some chains of the binary and ternary complexes,

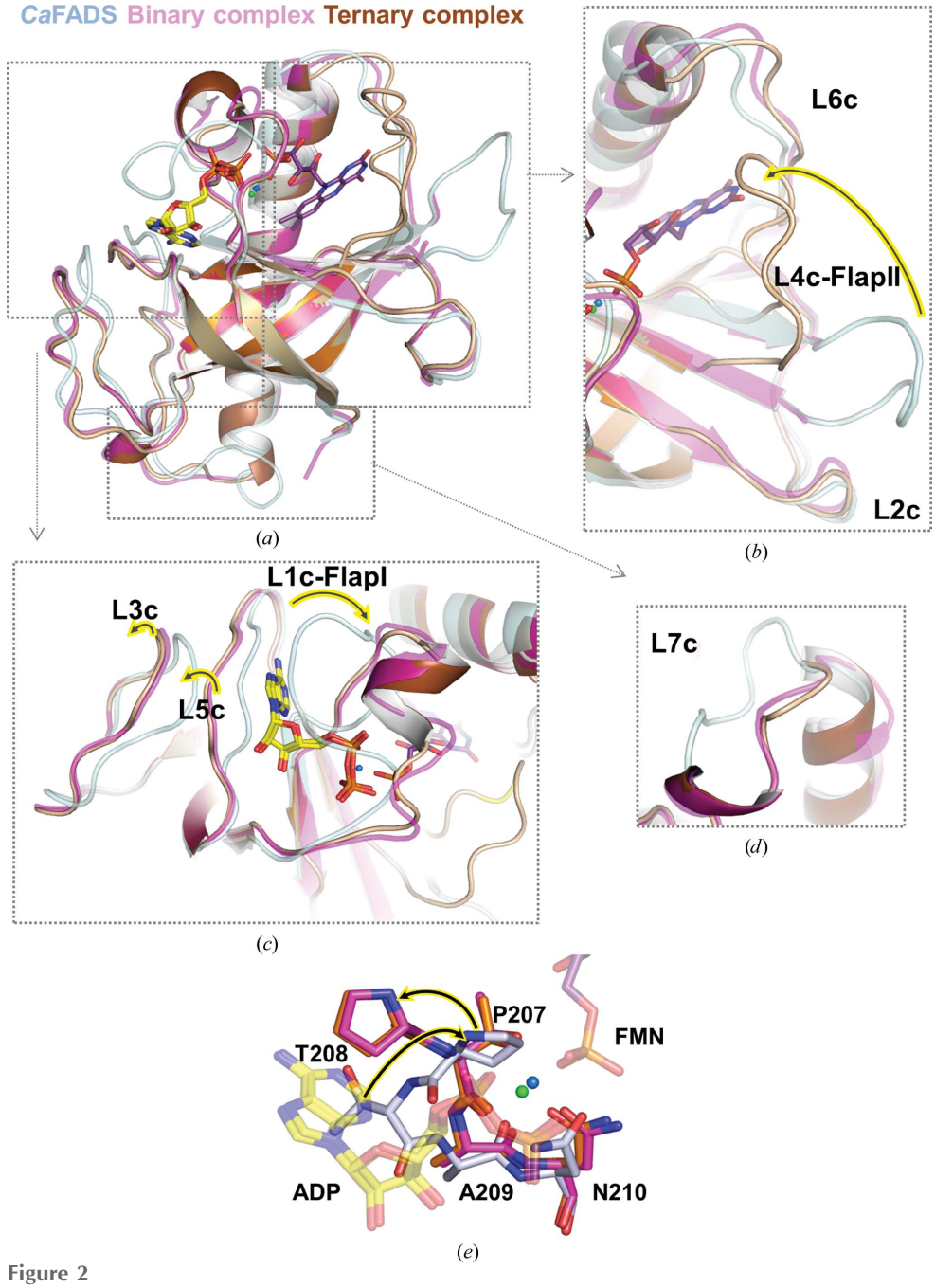


Figure 2
Superposition of the three available structures of the RFK module of *Ca*FADS: the free protein (light blue; PDB entry 2x0k), the ternary complex (orange-brown; PDB entry 5a89) and the binary complex (pink; PDB entry 5a88). (*a*) Overall view of the alignment of the RFK module, with the main differences between models highlighted with squares. (*b*) Zoom into the flavin-binding site, showing the displacement of L4c-FlapII and L6c in the ternary complexes. The L2c loop, which is involved in the interaction with the FMNAT module, does not show changes between the three structures. (*c*) Zoom into the adenine nucleotide-binding site, showing the three elements that change with respect to the wild type: L1c-FlapI, L5c and L3c. (*d*) Zoom into the C-terminus, showing the single conformation of L7c in the two complexes. (*e*) Conformation of the PTAN motif in the three models, showing the structural conformation changes of Pro207 and Thr208.

they are missing in others. Therefore, they do not affect the conformation of the loops, which is identical in each type of complex. Superposition with the apo structure of *Ca*FADS results in an r.m.s.d. on C^{α} atoms of just above 1 Å. However, there are considerable changes located around the two binding sites (Figs. 2*a*, 2*b* and 2*c*). In the ADP-binding region the concerted movement of loops L1c-FlapI, L3c and L5c results in opening of the ADP-binding site (Fig. 2*c*). In both complexes, the PTAN motif shows a dramatic reorganization (Fig. 2*e*), with Thr208 occupying the position of Pro207 in the apo structure and therefore releasing the adenine-ring binding site. Moreover, L1c-FlapI forms a short 3_{10}-helix-like structure that separates the flavin-binding and adenine-binding sites. In the ternary complex, conformational changes in L4c-FlapII and L6c result in closure of the FMN-binding site (Fig. 2*b*). L7c, which is located at the C-terminus after the long α1c helix, adopts only one of the two alternative conformations that were observed in free *Ca*FADS (Fig. 2*d*). Interestingly, the only loop that maintains its conformation with respect to free *Ca*FADS is loop L2c, which interacts with the FMNAT module in the full-length protein (Fig. 2*b*). This fact suggests that conformational changes in the RFK module triggered by ligand binding would not affect the FMNAT module in the monomeric full-length protein. This result is also in agreement with the ability of the RFK module to catalyse the phosphorylation of RF even in the absence of the FMNAT module.

Despite the similarity of the overall topology of the ternary complex presented here to those reported for the homologous monofunctional *Sp*RFK and *Hs*RFK (Supplementary Figs. S2 and S3; r.m.s.d. on C^{α} atoms of above 1.4 Å; Karthikeyan, Zhou, Osterman *et al.*, 2003; Bauer *et al.*, 2003), some of the loops connecting the secondary-structure elements show significant differences. This is the case for L3c, which shows an insertion of about 12 residues that are only observed in the enzymes from *Corynebacterium* spp. and *Mycobacterium* spp., as well as for the long loop L7c, which has no equivalent in *Hs*RFK or *Sp*RFK, where it is replaced by a C-terminal helical segment. Nevertheless, FlapI and FlapII show similar organizations in the bacterial and eukaryotic ternary complexes, providing an open but compact conformation for the adenine nucleotide-binding site and a closed and compact conformation for the flavin-binding pocket (Supplementary Figs. S3*a* and S3*b*). Also of interest, both ternary complexes show an identical arrangement of the conserved 207-PTAN-210 motif (Supplementary Fig. S3*c*), contrary to that observed for the apo form of *Ca*FADS.

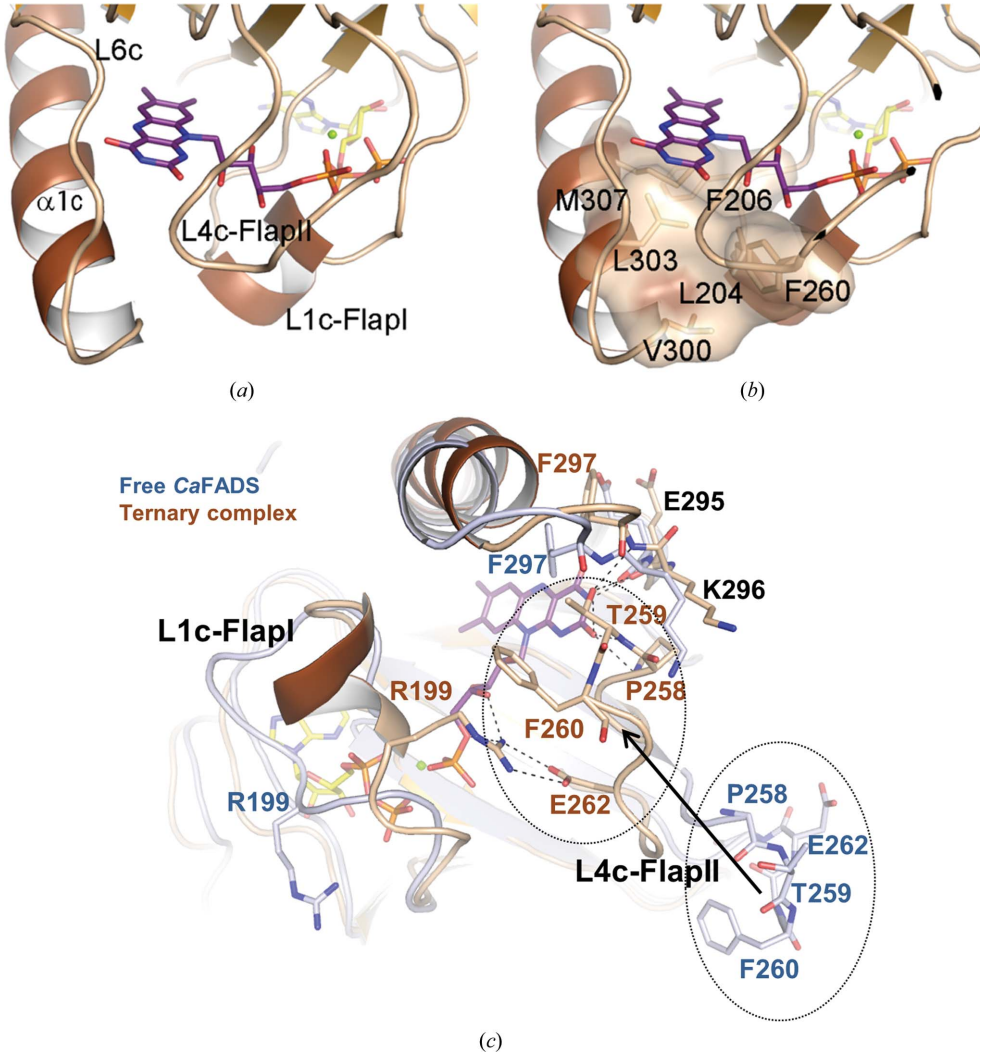


Figure 3
The flavin-binding site in the RFK module of *Ca*FADS according to the ternary complex. (*a*) Cartoon representation indicating the three main regions involved in the interaction of FMN: L1c-FlapI, L4c-FlapII and α1c–L6c. (*b*) Detail of the hydrophobic residues located at the *re*-face of the flavin ring. (*c*) Superposition of the free protein (light blue) and the ternary complex (brown), showing the residues interacting with FMN. The displacement of loop L4c-FlapII by more than 15 Å to interact with FMN is highlighted.

3.2. The flavin-binding site

In the ternary complex the FMN molecule is stabilized by hydrophobic interactions and by a network of hydrogen bonds involving three regions of the protein: (i) helix α1c and loop L6c, (ii) L4c-FlapII and L1c-FlapI and (iii) some residues located in the β-strands (the PTAN motif, Ser253, the catalytic base Glu268 and Phe270) (Fig. 3*a*, Supplementary Fig. S4). Various residues from L1c-FlapI (Leu204 and Phe206), L4c-FlapII

(Phe260) and α1c (Val300, Leu303 and Met307) form a hydrophobic core at the *re*-face of the isoalloxazine ring (Fig. 3*b*, Supplementary Fig. S4), whereas Val224, Met294 and Ser253 flank its *si*-face (Supplementary Fig. S4). Additionally, the isoalloxazine ring is hydrogen-bonded by Arg292, Glu295, Pro258 and Thr259 (Supplementary Fig. S4). It is noteworthy that Pro258, Thr259, Phe260 and Glu262 (all in L4c-FlapII) are displaced ~15 Å with respect to the apoprotein (Fig. 3*c*). The ribityl chain and phosphate group of FMN also establish hydrogen bonds to Asn210 from the PTAN motif and the catalytic base Glu268, as well as to Ala197, Gly198, Arg199 and Gly200 from L1c-FlapI (Supplementary Fig. S4).

The displacement of L4c-FlapII towards the flavin-binding site makes the flavin-binding cavity solvent-inaccessible (Figs. 3*c* and 4*b*). As mentioned above, some residues of this loop form hydrogen bonds to FMN or are part of the hydrophobic core in the *re*-face of the flavin ring. Moreover, L4c-FlapII interacts with the other regions involved in flavin stabilization: Thr259 forms hydrogen bonds to residues in the α1c helix (Glu295 and Phe297), whereas Asp262 and Phe260 form a salt bridge and a cation–π interaction with Arg199 (from L1c-FlapI). In the free structure L4c-FlapII presents high *B* factors and poor electron density, whereas in the binary complex it is completely absent, suggesting that this loop is flexible and adopts multiple conformations in the absence of the flavin. In summary, the binding of FMN induces a deep reorganization of the RFK module of *Ca*FADS, stabilizing a closed conformation of the flavin cavity by bringing together protein regions that are distant from each other in its absence.

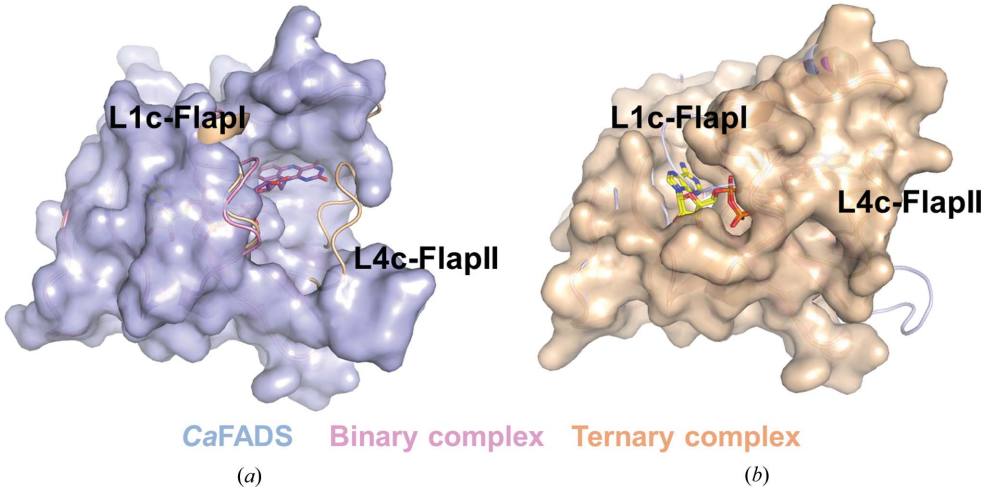


Figure 4
Solvent accessibility at the flavin-binding and adenine-binding sites of the RFK module of *Ca*FADS. (*a*) Superposition of the free *Ca*FADS RFK model and the binary and ternary complexes shown as cartoons with the free protein represented as a blue surface. The flavin-binding site is solvent-accessible in the free structure, whereas the adenine-binding site is closed. (*b*) Superposition of the same three structures in cartoon with the ternary complex represented as a brown surface. The adenine-binding site is open in both complexes, whereas the flavin-binding site is closed in the ternary complex.

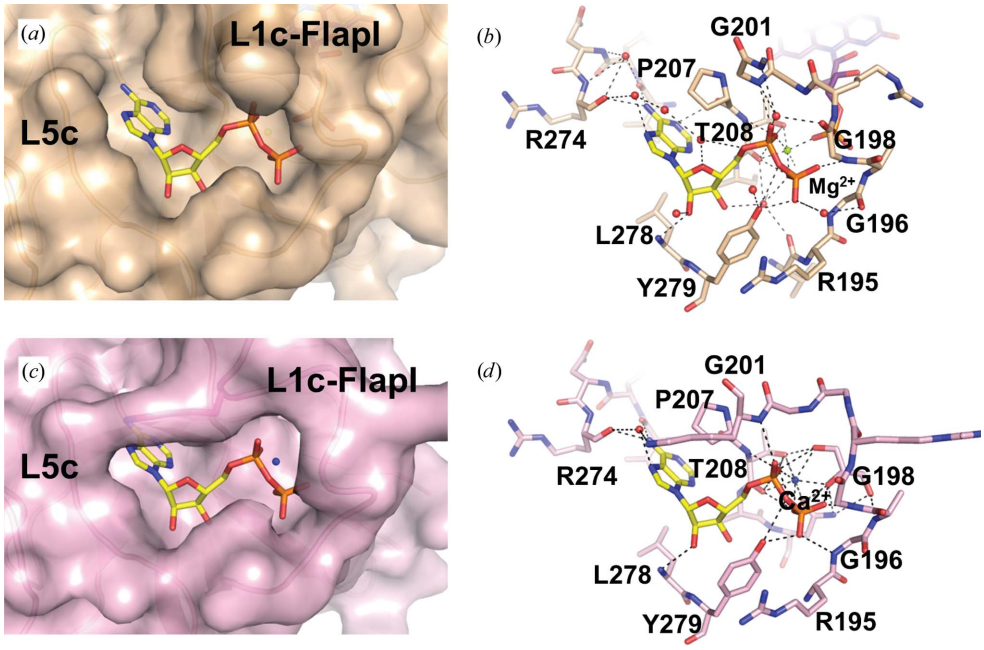


Figure 5
The adenine-binding site in the RFK module of *Ca*FADS. (*a*) Surface representation of the adenine-binding site in the ternary complex. (*b*) Detail of the residues interacting with ADP and Mg^{2+} in the ternary complex. (*c*) Surface representation of the adenine-binding site in the binary complex. (*d*) Detail of the residues interacting with ADP and Ca^{2+} in the binary complex.

3.3. The adenine-binding site

3.3.1. ADP-Mg^{2+} binding in the ternary complex. The electron density corresponding to the ADP molecule and Mg^{2+} ion is well defined in the ternary complex (Supplementary Fig. S5*a*). The ADP/ATP-binding pocket, which is closed in the ligand-free structure, is accessible to the solvent in the ternary complex owing to the concerted displacement of L1c-FlapI, L3c and L5c (Figs. 2*c* and 5*a*). The ADP molecule establishes hydrogen bonds to residues located in L1c-FlapI: Gly196, Gly198 and Gly201 interact with the α- and β-phosphates of ADP, whereas Pro207 and Thr208 (both belonging to the PTAN motif) form several hydrogen bonds *via* side-chain and main-chain atoms (Fig. 5*b*). Additional hydrogen bonds are observed between

ADP and residues from L5c, Leu271, Arg274, Leu278 and Tyr279, the latter being rotated with respect to the apo structure and forming a cation–π interaction with Arg195 (Fig. 5*b*). Mg^{2+} shows the typical octahedral coordination, interacting with three O atoms of the phosphate groups of FMN and ADP, with the side-chain O atom of Asn210 and with the main-chain and side-chain O atoms of Thr208 (Supplementary Fig. S5*b*).

3.3.2. ADP-Ca^{2+} binding in the binary complex. The density for ADP and Ca^{2+} is well defined in the four chains of the asymmetric unit (Supplementary Fig. 5*c*). The binding pocket is solvent-accessible, exhibiting a similar conformation to that of the ternary complex (Fig. 5*a*), with a slightly different conformation in L1c-FlapI (Fig. 5*c*). As mentioned before, chains *A* and *B* present low density for residues in the L1c-FlapI region. This indicates that L1c-FlapI is more flexible than in the ternary complex, where it is well defined. These data indicate that the ternary complex stabilizes L1c-FlapI. No other differences in the interactions between the complexes are observed (Figs. 5*b* and 5*d*). The Ca^{2+} ion (which replaces Mg^{2+}) is coordinated by the O atoms of ADP, Asn210 and Thr208 (Supplementary Fig. S5*d*). The sixth coordinating group is an O atom from a glycerol molecule (contained in the cryoprotection solution at a 20% concentration). The position of this molecule is similar to the phosphate group of FMN in the ternary complex (Supplementary Figs. S5*b* and S5*d*).

3.4. A new ion-binding site

The ternary complex contains an extra density corresponding to an ion tetrahedrally coordinated by two histidines (His186 and His325), Asp321 and a water molecule (Supplementary Fig. S6). This ion coordination with amino groups as donors and coordination number four is frequent for Zn^{2+} ions, whereas Mg^{2+} ions prefer an octahedral geometry with 'hard' oxygen ligands (Dudev & Lim, 2003). The $F_o - F_c$ maps calculated for both ions showed negative peaks for Mg^{2+} and positive peaks for Zn^{2+}, suggesting the presence of Zn^{2+} with partial occupancy, a fact confirmed by anomalous dispersion (Supplementary Fig. S6*d*). The Zn^{2+} ion is located outside the β-barrel at a distance of 24 Å from the active site. His186 is the first residue observed in the electron density (RFK comprises residues 183–338), while His325 and Asp321 are located at the C-termini of α1c and L7c, respectively (Supplementary Figs. S6*a* and S6*b*). Since this ion is far away from the active site, it is not likely to have any role in catalysis; rather, it seems to stabilize the conformation of the L7c and the N-terminus in the crystal of the independently produced RFK module. In the binary complex, where density for Zn^{2+} is not detected, L7c is completely disordered in chains *B* and *C*, while in chains *A* and *D* it presents high *B* factors. Observation of this ion only in the ternary-complex structure suggests its appearance by serendipity during purification or crystallization.

3.5. Catalytic properties of Δ(1–182)*Ca*FADS

Analysis of the formation of the products of the RFK activity confirmed that the isolated RFK module of *Ca*FADS catalyses the phosphorylation of RF. Moreover, as also reported for the full-length bifunctional protein, it exhibits strong RF-dependent inhibition profiles at saturating ATP concentrations (Supplementary Fig. S7; Serrano *et al.*, 2012; Serrano, Frago *et al.*, 2013). A fit of the data at different RF concentrations to the equation describing dead-end inhibition by excess RF generated values for the inhibition constant (K_i), $^{app}K_m^{RF}$ and $^{app}k_{cat}$ of 3.9 ± 1.0 µ*M*, 2.3 ± 0.7 µ*M* and 84 ± 24 min^{-1}, respectively, for Δ(1–182)*Ca*FADS. When the reaction was assayed at different ATP concentrations while keeping the RF concentration fixed at maximal activity before inhibition, the data fitted to a Michaelis–Menten model providing an $^{app}K_m^{ATP}$ and $^{app}k_{cat}$ of 17.0 ± 1.5 µ*M* and 25.8 ± 0.6 min^{-1}, respectively. In general, these values were within a factor of two/three of those reported in the case of the full-length enzyme (Serrano *et al.*, 2012; Serrano, Frago *et al.*, 2013). Therefore, in the absence of the FMNAT module the independently expressed RFK module of *Ca*FADS retains the ability to catalyse the phosphorylation of RF and conserves the structural determinants of the mechanism of inhibition by excess RF.

3.6. Binding of substrates to the RFK module of *Ca*FADS

The difficulties in obtaining crystal structures of the *Ca*FADS RFK module in complex with its substrates (ATP and RF) led us to use ITC to obtain some insight into the sequence of conformational changes produced upon binding of substrates. Under the assayed experimental conditions (PIPES and the absence of Mg^{2+}), both RF and ATP are able to bind to Δ(1–182)*Ca*FADS with slightly favourable enthalpic and entropic contributions (Fig. 6*a*), as previously described (phosphate and the absence of Mg^{2+}; Frago *et al.*, 2009). To determine whether an ordered binding applies, two additional experiments were performed: (i) a protein sample was pre-mixed with saturating concentrations of ATP (300 µ*M* ATP) and titrated with RF and (ii) a protein sample was pre-mixed with 180 µ*M* RF (solubility limit) and titrated with ATP. In the first case, only signals corresponding to nonspecific heat effects that were not related to a binding process were detected (not shown), indicating that RF does not bind to the preformed RFK–ATP complex. In the second case, titration of the preformed RFK–RF complex with ATP revealed an exothermic reaction (Fig. 6*b*), indicating RF binding prior to ATP to form the ternary substrate complex. In addition, the formation of this complex was driven by a large enthalpic contribution with a large opposing entropic contribution to the binding (Fig. 6*a*). This latter observation strongly indicates the formation of a large number of interactions as well as a large-scale conformational adaptation of the active site upon binding of the second substrate. Therefore, considering the apo and the product ternary-complex crystal structures, a large movement in the loop closing the flavin-binding site and a smaller but substantial one closing the ATP-binding site might be also expected when the enzyme changes from the substrate-free to the presumably more ordered and compact conformation of the ternary enzyme–substrate complex.

3.7. Simulation of Δ(1–182)*Ca*FADS dynamics

MD simulations were used to evaluate the coordination of conformational changes taking place in the RFK module of *Ca*FADS. Initial MD simulations of the free RFK structure showed high mobility for L4c-FlapII, but despite some tendency to move towards the α1c helix it did not close the RF-binding site (data not shown). During the simulation time L1c-FlapI was able to induce opening/closing of the ADP/ATP-binding site, as deduced from the increase in the distance between the centre of mass (W-M) of Arg199–Gly201 (in L1c-FlapI) and Asp275–Asp277 (in L5c) [W-M(199–201|275–277)] (Figs. 7*a*, 7*b* and 7*c*). Nevertheless, formation of the short 3_{10}-helix-like structure (present in both complexes; Fig. 2*c*) was not detected.

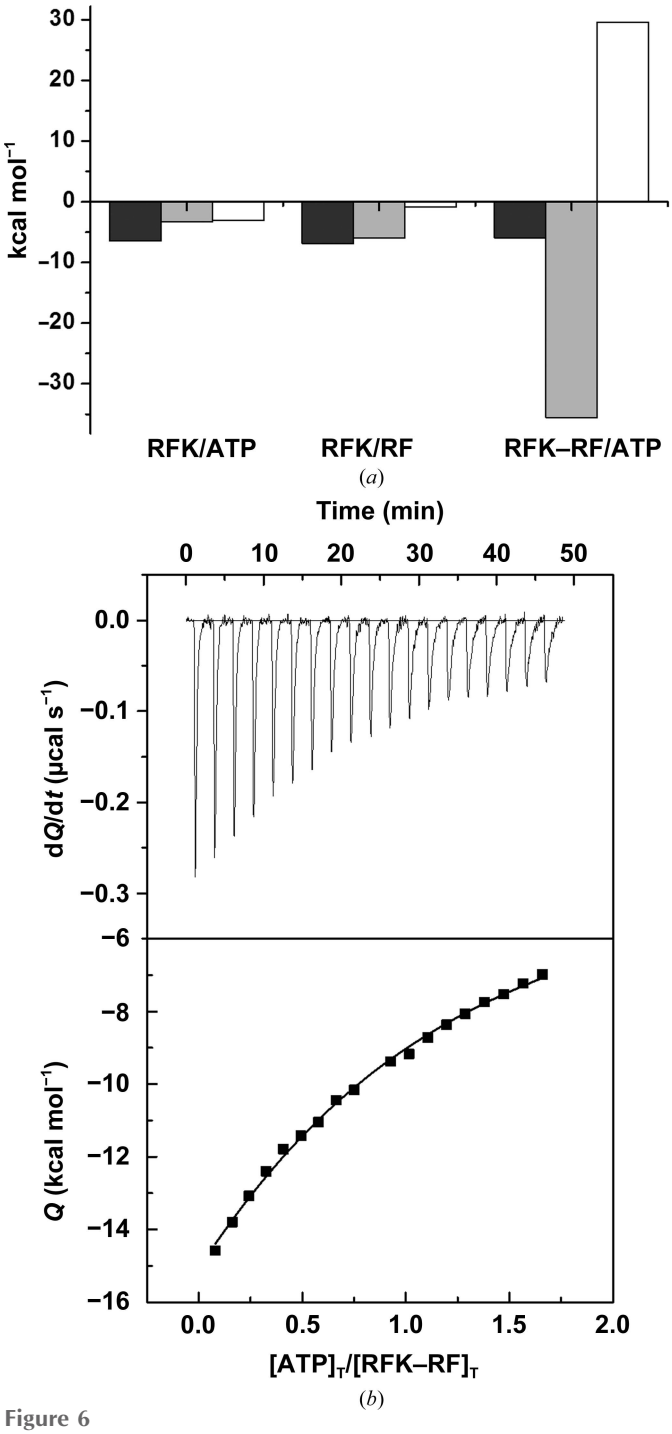


Figure 6
Thermodynamic dissection of the interaction of Δ(1–182)*Ca*FADS with its substrates. (*a*) Thermodynamic parameters for the titration of the RFK domain with either ATP or RF and of the preformed RFK–RF complex with ATP. The binding Gibbs energy (Δ*G*), enthalpy (Δ*H*) and entropy (−*T*Δ*S*) are represented as dark grey, light grey and white bars, respectively. (*b*) Calorimetric titration of the preformed RFK–RF complex with ATP. The upper panel shows the thermogram and the lower panel the corresponding isotherms with the integrated heat. Data were obtained at 25°C in 20 m*M* PIPES pH 7.0.

We built a second model by introducing either RF or FMN into the free RFK structure in the same orientation that they have in the ternary complex (referred to as the free RFK/RF and free RFK/FMN models, respectively). MD simulations of these models also showed high mobility for L4c-FlapII without closing of the flavin-binding site, as well as high mobility for L1c-FlapI with increasing frequency of opening of the ADP/ATP-binding site without formation of the short 3_{10}-helix-like structure (not shown) (Fig. 7*c*).

Our third model (referred to here as free RFK-MD/ADP) was built by adding ADP (with ternary-complex conformation and orientation) to the MD-equilibrated free RFK structure with a partially open ADP/ATP-binding site (free RFK-MD; Fig. 7*b*). The behaviour of this model during the MD simulation was similar to the other two models with regard to the ADP/ATP- and RF-binding sites and the short 3_{10}-helix-like structure (Fig. 7*c*). Interestingly, none of the above-discussed simulations showed conformational changes in the 206–210 consensus FPTAN motif with regard to the free RFK structure (Figs. 2*e* and 7*d*).

To further investigate when conformational changes in the PTAN motif take place, as well as the regions that contribute to these changes, we used targeted MD simulations. In these simulations the starting points were either the free RFK structure or some of the models obtained in the above simulations. Target conformations of particular regions were fixed to those in the crystal structure of the ternary complex. Initially, we targeted the formation of the short 3_{10}-helix-like structure in L1c-FlapI. This simulation produced opening of the ADP/ATP-binding site and partial displacement of the Phe206-Pro207 peptide bond towards its conformation in the binary- and ternary-complex crystals, but the conformation of Thr208 did not change. We only observed the expected Phe206-Pro207 peptide-bond conformational change when, using the same targeted structure, ADP and FMN were placed in the free RFK-MD model in the ternary-complex crystal conformation (referred to here as free RFK-MD/ADP/FMN; Fig. 8). Therefore, neither the Phe206-Pro207 peptide-bond conformational change upon L1c-FlapI displacement nor the formation of the 3_{10}-helix-like structure nor the presence of the adenine nucleotide by themselves appear to be key to inducing the conformational change of the Thr208 backbone.

These observations leave the possibility of the Thr208 backbone conformational change relying on the coordination of the Thr208 side chain by the adenine nucleotide *via* the metal ion Mg^{2+} (not explicitly defined in the force field). To

initially check for such a possibility, the free RFK structure was phosphorylated on the Thr208 side chain (free RFK-ThrP). MD simulation on this model resulted in very fast formation of the L1c-FlapI 3_{10}-helix-like structure and in the opening of the ATP-binding site (Figs. 9*a*, 9*b* and 9*c*). However, the conformation of the FPTAN motif did not change.

A final model was constructed by placing ADP-Mg^{2+} in the Thr208 environment of the free RFK-ThrP equilibrated model to mimic coordination between Mg^{2+} and the side-chain O atom of Thr208 (free RFK-MD/ADP/Mg^{2+}). After equilibration, we carried out a targeted MD simulation upon the conformation of ADP and the residues forming the FlapI 3_{10}-helix-like structure. This simulation produced the expected conformational change in the 206–210 (FPTAN) backbone (Figs. 9*d* and 9*e*). These results make the coordination of Thr208 with Mg^{2+} upon ATP-Mg^{2+} binding feasible, leading to the 206–208 backbone conformational change and enabling

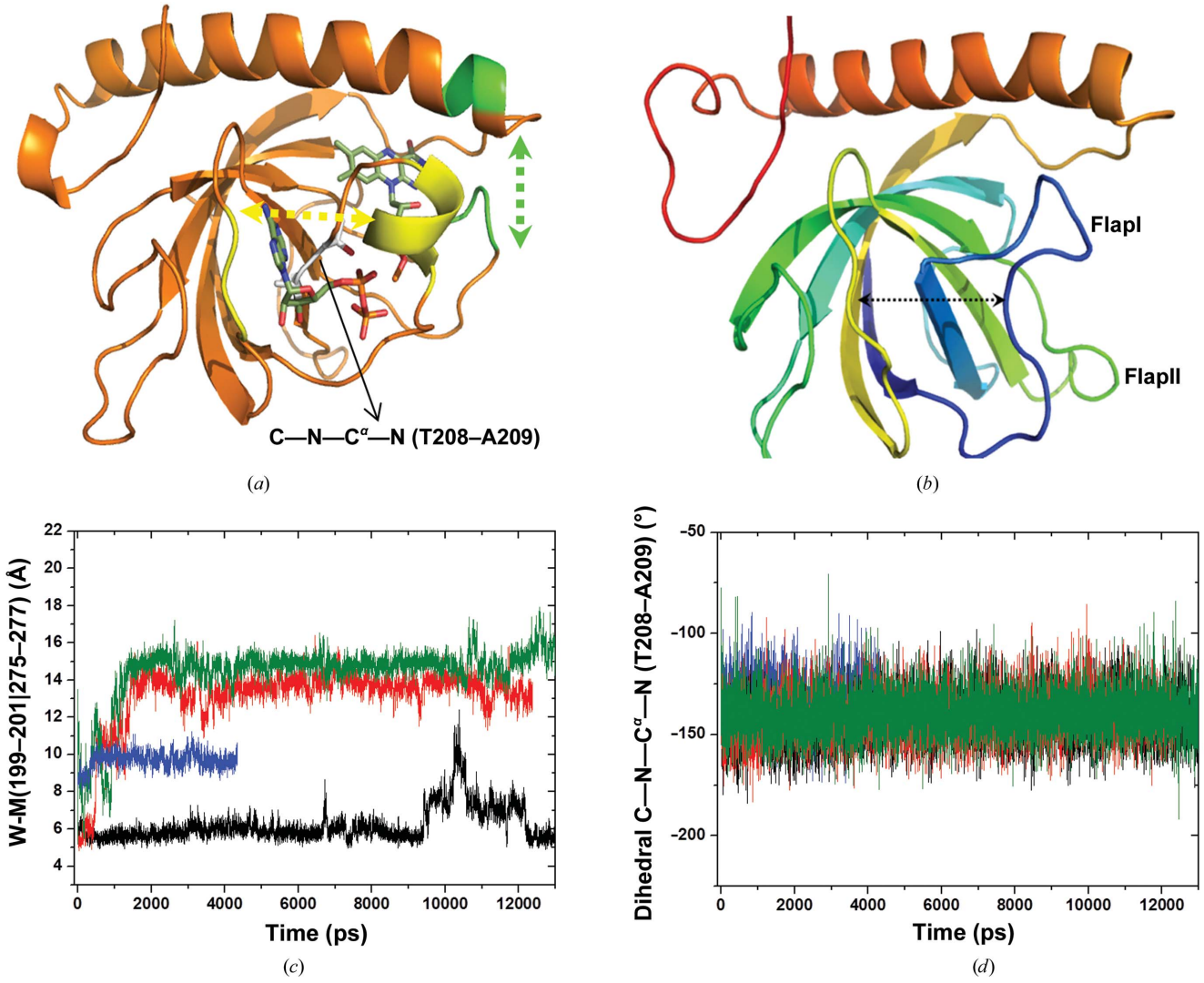


Figure 7
MD simulations of the RFK module of free *Ca*FADS in the presence of flavins and in the presence of ADP. (*a*) Structure of the ternary complex showing the conformational elements used in the MD evaluation. The distance between the centre of mass of residues Pro258–Phe260 (in L4c-FlapII) and Glu301–Leu303 (in α1c) [W-M(258–260|301–303), green atoms and line] was taken as an indicator of the opening/closing of the RF-binding site, while the distance between the centre of mass of residues Arg199–Gly201 (in L1c-FlapI) and Asp275–Asp277 (in L5c) [W-M(199–201|275–277), yellow atoms and line] was taken as an indicator of the opening/closing of the ADP/ATP-binding site. The dihedral angle C—N—C^{α}—N between residues Thr208–Ala209 [dihedral C—N—C^{α}—N (T208–A209), grey atoms in sticks] was taken as an estimation of the conformation of the FPTAN consensus motif: this angle is −150.38 and −77.36° in the crystal structures of free RFK and the ternary complex, respectively. (*b*) Snapshot at 10.385 ns of a simulation of the free RFK model: it shows displacement of FlapI and opening/closing of the ADP/ATP-binding site. (*c*) Trajectory of the W-M(199–201|275–277) distance during MD of the free RFK module (PDB entry 2x0k chain *A*, black), free RFK/RF (red), free RFK/FMN (green) and free RFK-MD/ADP (blue) models. The free RFK/RF and free RFK/FMN models were built by introducing RF and FMN, respectively, into the free RFK structure in the conformation of FMN in the ternary-complex crystal. The free RFK-MD/ADP model was built by introducing the ADP into an equilibrium structure obtained during the MD simulation of the free RFK model. (*d*) Trajectory of the dihedral angle C—N—C^{α}—N (T208–A209) during the simulation of the free RFK (PDB entry 2x0k chain *A*, black), free RFK/RF (red), free RFK/FMN (green) and free RFK-MD/ADP (blue) models. The first points in the plots in (*c*) and (*d*) are the corresponding values in the crystal ternary complex.

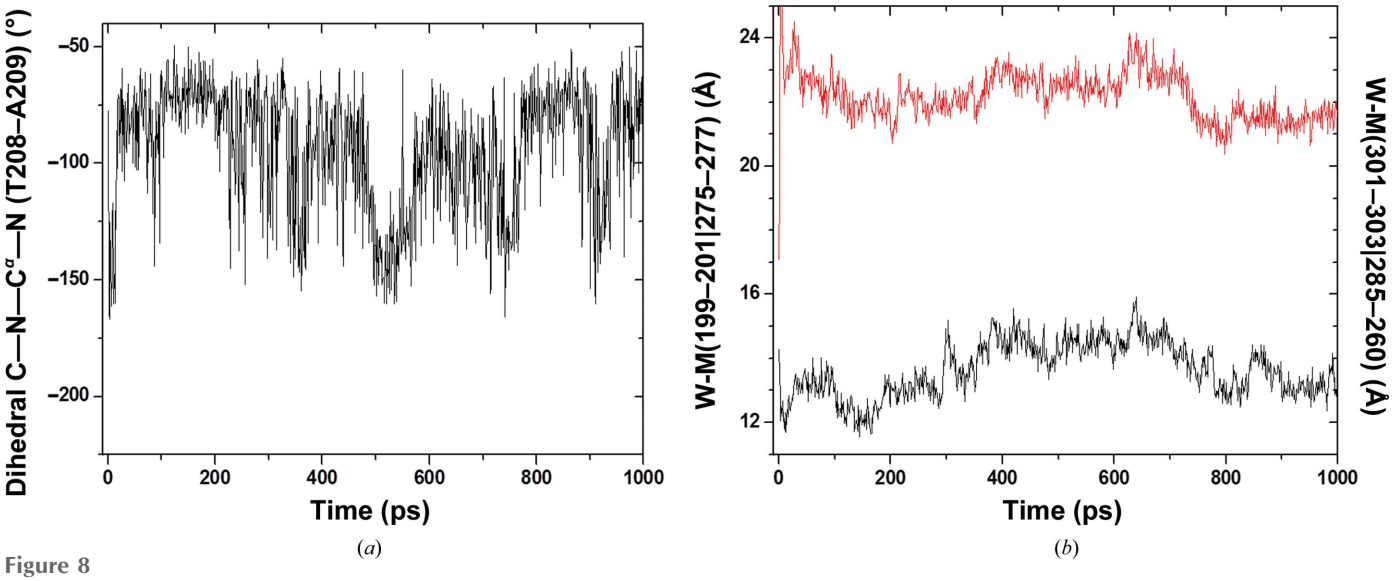


Figure 8
MD simulations in the ternary complex: conformational changes at the FPTAN motif and opening/closing of the ADP/ATP- and RF/FMN-binding sites in the targeted MD simulation of the model upon ADP, FMN and the residues forming the 3_{10}-helix-like structure in the ternary crystal structure (L1c-Flap I). (*a*) Trajectory of the dihedral angle C—N—C^{α}—N (T208–A209). (*b*) Trajectories of the W-M(199–201|275–277) (black) and W-M(301–303|258–260) (red) distances. The first point in each plot corresponds to the value in the structure of the ternary-complex crystal. The ternary-complex model (free RFK-MD/ADP/FMN) was built from the MD free RFK model with the adenine nucleotide-binding site partially open (free RFK-MD) by simultaneously adding ADP and FMN in the conformation found in the ternary-complex crystal.

Thr208 to interact with the ribose of the RF substrate (or with the phosphate of the FMN product in the ternary-complex crystal structure).

4. Discussion

Ternary and binary complexes of the RFK module of *Ca*FADS with the products of its reaction, FMN and ADP-Mg^{2+}, exhibit significant structural rearrangements with respect to the free apoprotein, affecting the conformation of the consensus PTAN motif as well as of most of the protein loops. The exception is L2c, which is involved in the interaction between the RFK and FMNAT modules in the *Ca*FADS protomer, this being additional evidence of the independence of the two modules within one protomer (Herguedas *et al.*, 2010). Binding studies in solution also indicate important conformational changes upon binding of the substrates of the RFK reaction, showing that although RF binds prior to ATP, the closed conformation of the ternary-substrate complex will be attained upon binding of the second substrate. Besides, together with the available structural and enzymatic data, it can be confirmed that the structural changes triggered by ligand binding (or catalysis) in the RFK module cannot be transferred to the FMNAT module within the same protomer. However, as will be analysed later, structural changes in the RFK module must also influence the FMNAT active site of the neighbouring protomer in the apo structure, as previously reported for the dimer of trimers of *Ca*FADS, as the loops that show the most significant changes are located in the interface between protomers of each trimer. Some of the conformational changes observed in the ternary complex (such as the displacement of L4c-FlapII) have previously been reported for *Hs*RFK, whereas the changes in L3c, L5c and especially in L1c-FlapI and in the PTAN motif appear to be specific to the bifunctional *Ca*FADS. In the binary complex, structural rearrangements mainly affect the ADP-binding site, which exhibits a similar conformation to the ternary complex, whereas L4c-FlapII is not included in the final model as it presented poor electron density.

4.1. Ligand binding induces structural rearrangements in L1c-FlapI, L3c, L4c-FlapII and L5c

In *Sp*RFK, FlapI (which contains several conserved glycines involved in the interaction with the phosphates of ADP in the ternary complex) separates the flavin- and the adenine nucleotide-binding sites and makes both binding sites accessible in the apo structure (Bauer *et al.*, 2003; Karthikeyan, Zhou, Mseeh *et al.*, 2003; Karthikeyan, Zhou, Osterman *et al.*, 2003; Supplementary Fig. S2*b*). On the contrary, in the apo RFK module of the *Ca*FADS crystal structure the adenine-binding site was closed owing to the conformation of L3c, L5c and L1c-FlapI, which interact with each other and close the cavity (Figs. 2*c*, 4*a* and Supplementary Fig. S2*a*). The formation of the ternary-product complex induces a conformational change in the L1c-FlapI region, which is displaced towards α1c, with part of it forming a short 3_{10}-helix-like structure. The loss of interactions with L1c-FlapI releases loops L5c and L3c, which in turn move in the opposite direction and expand the cavity (Figs. 2*c* and 4*b*). Since the binary complex shows the same relative arrangement in these three loops, these structural changes are produced by the binding of the adenine nucleotide without the need for simultaneous binding of RF/FMN. The lower density observed in L1c-FlapI in some protein chains in the binary complex also indicates that its conformation is more flexible than in the ternary complex. As

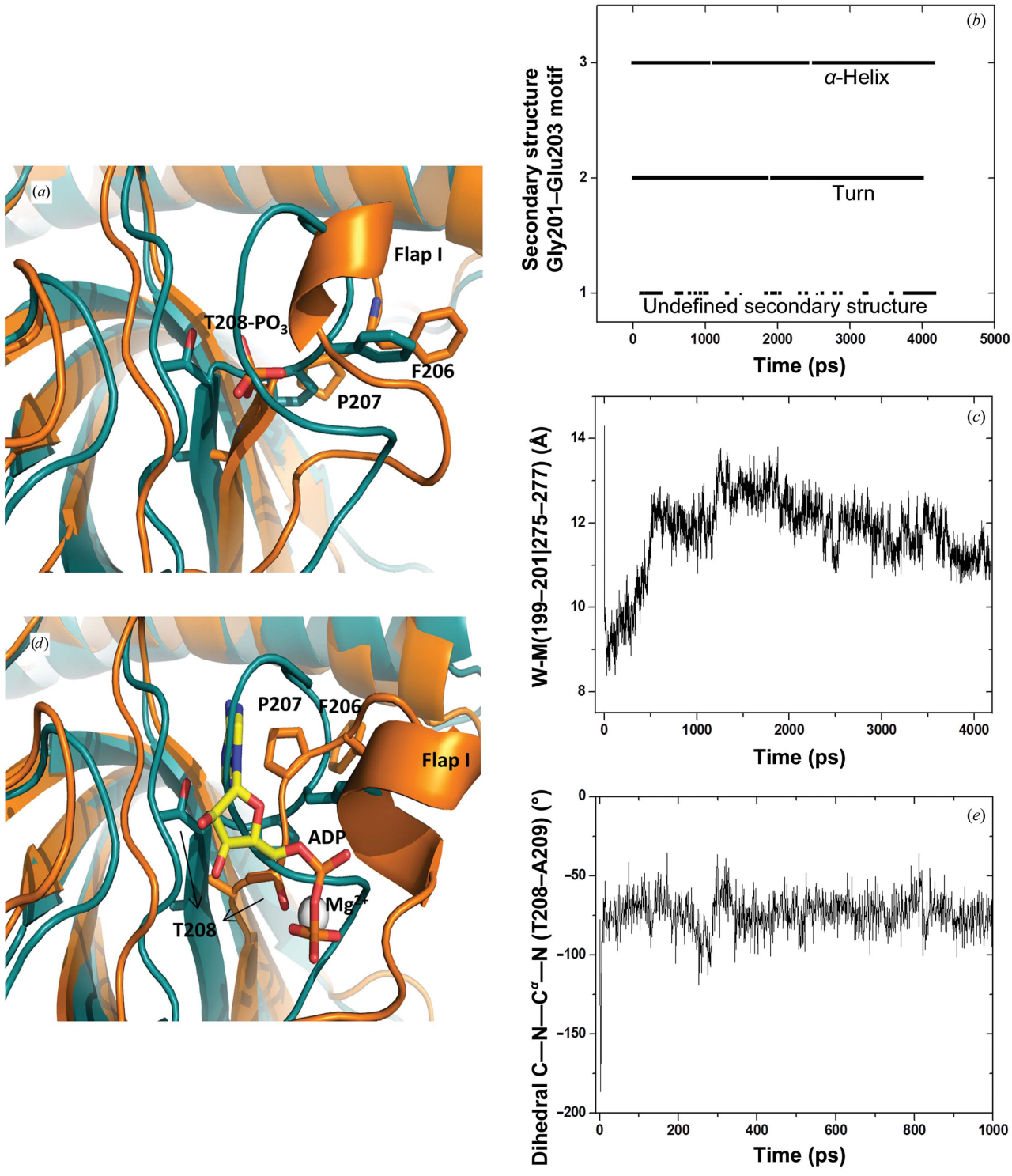


Figure 9
Coordination of the Thr208 side chain and conformational changes in the FPTAN motif. (*a*) MD simulation of the free RFK structure (PDB entry 2x0k chain *A*) with a phosphorylated Thr208 side chain (free RFK-ThrP model) resulted in very fast formation of the FlapI 3_{10}-helix-like structure. The structure after 3.664 ns of simulation (orange) is superimposed on that of the free RFK module (PDB entry 2x0k chain *A*, blue). (*b*) Secondary-structure assignments of residues Gly201–Glu203 during the MD simulation according to the *DSSP* method (Kabsch & Sander, 1983*a*,*b*). (*c*) Trajectory of the W-M(199–201|275–277) distance during the MD simulation. (*d*) Targeted MD simulation of a free RFK-ThrP-MD/ADP/Mg^{2+} model onto the conformation of ADP and the FlapI 3_{10}-helix-like structure in the ternary-complex crystal showed the predicted conformational change in the FPTAN motif. The structure after 684 ps of simulation (orange) is superimposed on that of the free RFK module (PDB entry 2x0k chain *A*, blue). (*e*) Trajectory of the dihedral angle C—N—C^{α}—N (T208–A209). The first points in (*c*) and (*d*) correspond to the respective values in the ternary-complex crystal structure. The free RFK-ThrP-MD/ADP/Mg^{2+} model was built using the equilibrated free RFK-ThrP model by placing ADP-Mg^{2+} in the Thr208 environment to mimic the coordination between Mg^{2+} and the hydroxyl group of Thr208.

several residues of L1c-FlapI hydrogen-bond to FMN in the ternary complex, the lack of this ligand in the binary complex explains the observed higher flexibility. MD simulations of the free RFK module suggest that in solution L3c, L5c and particularly L1c-FlapI are displaced and also populate some conformations that are compatible with an open–closed equilibrium for the adenine nucleotide-binding site (Fig. 7*c*). These simulations also indicate that the presence of either flavin or adenine nucleotides increases the opening probability of this site, even though they are not able to induce formation of the 3_{10}-helix-like structure during the simulation time. Therefore, simulations also show the closed apo RFK module structure being able to evolve to the active conformation and not being an autoinhibited conformation trapped in the crystal.

L4c-FlapII is disordered in the binary complex, with high *B* factors and poor electron density in the free form, resulting in a large open cavity for flavin binding (Fig. 4*a*). Binding of FMN in the ternary complex induces a change in the size, shape and solvent-accessibility of the cavity, which was partially occupied by the side chain of Phe270 in the apoprotein. Thus, FMN binding causes a displacement of more than 15 Å in L4c-FlapII towards the flavin-binding site (Figs. 1*b* and 4*b*). The new conformation of L4c-FlapII is stabilized by interactions with FMN, L1c-FlapI, the α1c helix and L6c, all of which also contribute to the closure of the cavity (Fig. 4*b*).

4.2. The consensus PTAN motif stabilizes ligand binding and promotes the formation of the catalytic competent complex

The consensus 207-PTAN-210 motif that is present in all of the RFK modules of bifunctional FADSs as well as in monofunctional eukaryotic RFKs is proposed to stabilize the position of the phosphates of ATP and the Mg^{2+} ion during catalysis (Karthikeyan, Zhou, Mseeh *et al.*, 2003; Karthikeyan, Zhou, Osterman *et al.*, 2003; Bauer *et al.*, 2003; Serrano, Frago *et al.*, 2013). The crystal structures presented here reveal that in *Ca*FADS this motif undergoes an extensive rearrangement after binding of the adenine nucleotide. In the binary and ternary complexes the adenine ring of ADP is situated at the position of the side chain of Thr208 in the apo enzyme, which is now displaced to the position previously occupied by Pro207 (Fig. 2*e*). These conformational changes result in opening of the ADP/ATP-binding cavity. Since in the ternary complex Thr208 is the only residue that interacts with FMN, ADP, Mg^{2+} and the Glu268 catalytic base, it appears to be essential for optimal ligand-binding and enzyme activity. This observation is supported by the characterization of the T208A and T208D mutants, which lack the ability to catalyse the RFK activity and to bind the substrates of the RFK reaction, ATP-Mg^{2+} and RF (Serrano, Frago *et al.*, 2013). Furthermore, Asn210 interacts with FMN, Mg^{2+} and Glu268. Mutations introduced at this residue also have drastic negative effects on the RFK activity, probably owing to binding of the substrates in a noncatalytic conformation (Serrano, Frago *et al.*, 2013). Therefore, the interactions of the Asn210 side chain with FMN, Mg^{2+} and Glu268 observed in the ternary complex appear to be critical to set the geometry of the catalytically competent complex, particularly contributing to maintaining the γ-phosphate of ATP donor and the OH acceptor of RF in the correct orientation. Thus, the final orientation observed in both the ternary and binary complexes for this motif, particularly for the Pro-Thr backbone, is similar to that found in *Hs*RFK and *Sp*RFK, either free or in complex. Our MD simulations indicate that just the presence of the flavin or adenine nucleotide is not sufficient to induce the conformational change of the Thr208 backbone (Fig. 7*d*). Thus, the coordination of the Thr208 side chain by the adenine nucleotide *via* the metal ion, Mg^{2+}, appears to be the most likely determinant for the 207-PTAN-210 motif to achieve the catalytic geometry.

4.3. Structural changes for FMN biosynthesis

Previous enzymatic characterizations of *Ca*FADS and rat RFK suggested that FMN is synthesized through an ordered bi-bi mechanism in which RF binds first into the active site followed by ATP and, after RF has been phosphorylated, ADP is released before FMN (Yamada *et al.*, 1990; Efimov *et al.*, 1998). The available crystal structures for the RFK module of *Ca*FADS represent the initial state (apo structure in the full-length protein) and two complexes with products of the reaction: one with both products, ADP-Mg^{2+} and FMN, and the other only with ADP-Mg^{2+}.

In the absence of ligands, and contrary to that found in other RFKs, the residues of the PTAN motif occupy the ATP/ADP-binding site and the access to the adenine nucleotide cavity is closed by the interaction of L1c-FlapI with L5c, although an open–closed equilibrium for this site is envisaged (Fig. 7*c*). In this state the protein shows a large cavity for RF/FMN binding without structural rearrangements. Studies in solution showed the ability of RF, although not of FMN, to bind Δ(1–182)*Ca*FADS in the absence of adenine nucleotides (Fig. 6*a*; Frago *et al.*, 2009). Nevertheless, considering that this crystal structure remains elusive and the low enthalpic and entropic contributions to the binding, such RF binding might not be catalytically productive.

In the binary complex, binding of ADP-Mg^{2+} induces structural changes in the PTAN motif and in L1c-FlapI, which adopts an arch-shaped structure and separates the two binding sites. Moreover, the conformation of L1c-FlapI, L3c and L5c is very similar in the binary and ternary complexes, suggesting that ADP-Mg^{2+} is able to induce the changes in the whole area, even in the absence of RF or FMN. These data are consistent with the characterization by isothermal titration calorimetry, indicating that both ATP and ADP can bind to the RFK module even in the absence of flavins (Fig. 6*a*; Frago *et al.*, 2009), as well as with an open-closed equilibrium for the adenine nucleotide-binding site in the apo structure (Figs. 7*b* and 7*c*).

In the product ternary complex, in addition to changes in the PTAN motif and in the adenine-binding site, flavin binding is accompanied by structural changes in L4c-FlapII, which is displaced towards the flavin-binding site. It is likely that given

the wide cavity in the free form and the number of interactions of the ribityl chain with the protein, the RF is not oriented properly until the adenine nucleotide, the PTAN motif, L1c-FlapI and L4c-FlapII have rearranged to adopt the closed flavin-binding site structure. These structural observations agree with the positive cooperativity of ATP/ADP-Mg^{2+} in flavin binding to Δ(1–182)*Ca*FADS and with the strong stabilization of the product ternary complex by modulating the thermodynamic parameters, which enhance the affinity of RF and FMN (Frago *et al.*, 2009). Crystal structures have not been obtained for this ternary complexes. Nevertheless, the binding studies in solution presented here indicate that although RF binding takes place prior to that of ATP, RF only attains its catalytic configuration in the binding cavity after binding of the second substrate (Fig. 6). These studies also predict large-scale conformational adaptation from the binary substrate complex to the ternary substrate complex compatible with the production of the presumably more ordered and compact conformation upon closing the RF-binding site.

In the catalytically competent complex, Glu268 and Asn210 play an essential role, with the first acting as the catalytic base and the second favouring the approach of RF and ATP-Mg^{2+} so that nucleophilic attack occurs. Currently, it is unknown whether the reaction takes place with a pentavalent intermediate or by forming a metaphosphate intermediate. In any case, a single basic residue interacts in this region and could stabilize the intermediate: Arg199. It is noteworthy that this residue also contributes to the stabilization of the closed conformation of the active site. Once transfer of the phosphoryl group occurs, ADP-Mg^{2+} may be released (its cavity is open), causing movement of L1c-FlapI (which would collapse onto L5c). As a result, L1c-FlapI breaks its interactions with L4c-FlapII, which in turns becomes more mobile, allowing FMN release. Altogether, these data are compatible with the hypothesis of ADP leaving the phosphorylation site prior to FMN (Efimov *et al.*, 1998).

Therefore, the structural, simulation and binding data presented here support binding of ATP-Mg^{2+} being required prior to RF attaining the catalytic geometry in the active site, in agreement with previous data, suggesting that release of the ADP product would take place prior to dissociation of FMN.

4.4. Structural changes in the RFK module and the full-length protein structure

Despite the massive conformational changes observed in the binary and ternary complexes, the loop involved in the interaction with the FMNAT module of the same protomer, L2c, maintains its conformation in the three states (Fig. 2*b*). This suggests that the structural changes in RFK could not be transferred to the FMNAT site of the same protomer. The kinetic analysis of the RFK module also shows that this

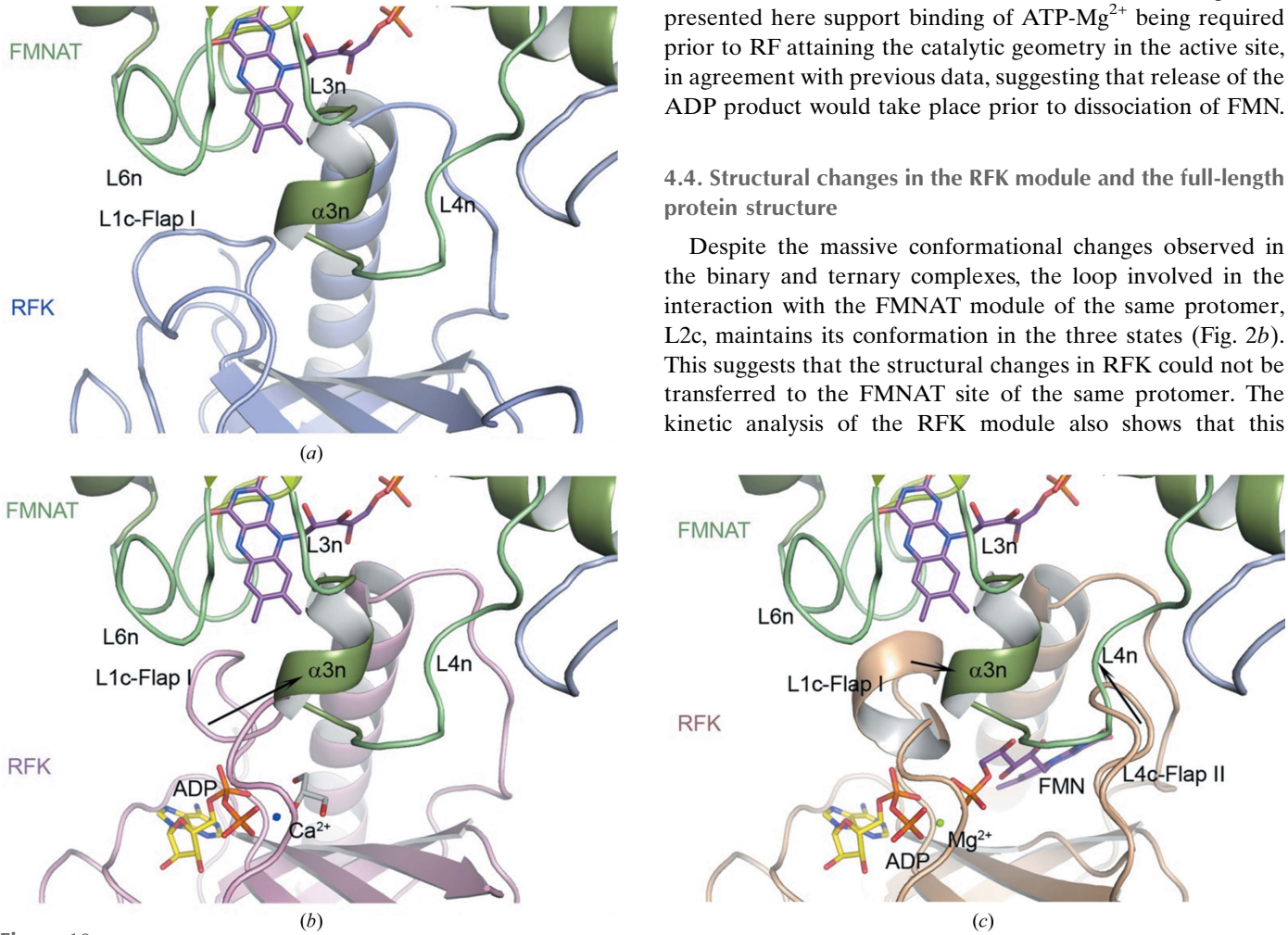


Figure 10
Influence of the conformational changes in the RFK module on the quaternary organization of *Ca*FADS. (*a*) Cartoon representation of the interface between the RFK and FMNAT modules of different chains in the hexameric protein, highlighting the main structural elements involved in the interactions. (*b*) Model of the interface between RFK and FMNAT in the presence of ADP-Mg^{+}. Conformational changes in L1c-FlapI, which forms a 3_{10}-helix-like structure, should modify the α3n region of the FMNAT module. (*c*) Model of the interface between RFK and FMNAT modules in the presence of FMN and ADP-Mg^{2+}. Conformational changes in L1c-FlapI and L4c-FlapII should induce changes in α3n and L4n. Arrows indicate expected clashes between secondary-structure elements.

module is active, although its kinetic parameters slightly differ from those of the full-length protein. These observations suggest that the conformational changes in the RFK module should be similar in the full-length protein and indicate that the FMNAT module also contributes to modulate the RFK activity in *Ca*FADS. Different studies with mutants have also shown that both modules interact functionally, as substitutions in the RFK module (Pro208, Thr210 and Glu268) have a negative effect on the FMNAT activity, while mutations in the FMNAT module also impact the RFK activity (Serrano, Frago *et al.*, 2013; Serrano *et al.*, 2012, 2015). This interplay between modules can only be explained in the context of the oligomeric structure. The dimer-of-trimers structure was first observed in *Ca*FADS crystals and subsequently by AFM as low-resolution topography images resembling this ensemble in shape and in volume (Marcuello *et al.*, 2013; Herguedas *et al.*, 2010; Frago *et al.*, 2007). This latter study also determined the monomer as the main transient form in solution, with oligomerization being triggered by the binding of substrates and/or products. Thus, AFM also showed that while binding of the ADP-Mg^{2+} product stabilized an amorphous dimer of trimers (failing to form the compact dimer of trimers), ATP-Mg^{2+} produced a compact dimer of trimers. Superposition of binary and ternary complexes on the quaternary apo *Ca*FADS structure shows that the conformational changes in the RFK module will impact the flavin-binding site at the FMNAT module, affecting L3n (residues 55–57), α3n (58–62) and L4n (63–73) (Fig. 10). The low resolution of the AFM data makes it impossible to distinguish atomistic details such as interaction surfaces between protomers or the conformation of loops at these surfaces. Thus, it is very likely that these differ from those observed in the apo quaternary organization, despite both corresponding to a dimer of trimers. In this context, we cannot speculate on the extent of these changes until structures containing the FMNAT module in complex with ligands become available. To date, it is not clear which factors will determine the monomer–oligomer equilibrium *in vivo*. Nevertheless, in the monomeric state, which is the main apo form in solution, the RFK and FMNAT modules might work independently, while in the oligomeric ensemble they will influence each other. Prokaryotic organisms have FMN- and FAD-containing flavoproteins and therefore the channelling of FMN in the hexamer will only be required for the production of FAD, whereas FMN will have to be released from the RFK module when FMN is required. Therefore, the crystal structures here presented reflect the final state of the kinase reaction, and further studies will be needed to understand the full catalytic cycle of *Ca*FADS.

So far, there is no experimental information about whether or not other prokaryotic FADSs might undergo quaternary organizations into dimers of trimers. Previous studies predicted L3c in the RFK module of *Ca*FADS to be a key determinant in the stabilization of the dimer of trimers (Marcuello *et al.*, 2013). This loop is only present in the *Corynebacterium* and *Mycobacterium* species, suggesting that stabilization into the dimer of trimers formed in *Ca*FADS during the catalytic cycle might be a particular feature of just a few FADSs. In this context, recent studies suggest that RFKs, FMNATs and FADSs could interact with different apo flavoproteins, providing them with their cofactors (Park *et al.*, 2012; Yazdanpanah *et al.*, 2009; Giancaspero *et al.*, 2015). This observation agrees with the fact that FADSs and RFKs are co-purified with tightly bound flavins, which can only be removed by using hydrophobic columns during purification. Therefore, the interaction with a broad spectrum of proteins will be an additional factor to consider when studying the structural bases of the full catalytic cycle, as the release of flavins could be modulated by the interaction with their apo flavoprotein counterparts.

Acknowledgements

This work has been supported by MINECO, Spain (BIO2013-42978-P to MM and BFU2014-59389-P to JAH), the Aragonian Government-FEDER (B18), Autonomous Community of Madrid (S2010/BMD-2457), Departamento Administrativo de Ciencia, Tecnología e Innovación (COLCIENCIAS) and Universidad Industrial de Santander (project 1818 to IL). The authors thank A. Velázquez-Campoy for helpful discussions in the interpretation of the ITC experiments.

References

Antony, J., Medvedev, D. M. & Stuchebrukhov, A. A. (2000). *J. Am. Chem. Soc.* **122**, 1057–1065.
Barile, M., Brizio, C., Valenti, D., De Virgilio, C. & Passarella, S. (2000). *Eur. J. Biochem.* **267**, 4888–4900.
Bas, D. C., Rogers, D. M. & Jensen, J. H. (2008). *Proteins*, **73**, 765–783.
Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Bauer, S., Kemter, K., Bacher, A., Huber, R., Fischer, M. & Steinbacher, S. (2003). *J. Mol. Biol.* **326**, 1463–1473.
Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B. & Woods, R. J. (2005). *J. Comput. Chem.* **26**, 1668–1688.
Cheek, S., Ginalski, K., Zhang, H. & Grishin, N. V. (2005). *BMC Struct. Biol.* **5**, 6.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W. & Kollman, P. A. (1995). *J. Am. Chem. Soc.* **117**, 5179–5197.
Delano, W. L. (2002). *PyMOL.* http://www.pymol.org.
Dudev, T. & Lim, C. (2003). *Chem. Rev.* **103**, 773–788.
Efimov, I., Kuusk, V., Zhang, X. & McIntire, W. S. (1998). *Biochemistry*, **37**, 9716–9723.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D**69**, 1204–1214.
Frago, S., Goñi, G., Herguedas, B., Peregrina, J. R., Serrano, A., Pérez-Dorado, I., Molina, R., Gómez-Moreno, C., Hermoso, J. A., Martínez-Júlvez, M., Mayhew, S. G. & Medina, M. (2007). *Arch. Biochem. Biophys.* **467**, 206–217.
Frago, S., Martínez-Júlvez, M., Serrano, A. & Medina, M. (2008). *BMC Microbiol.* **8**, 160.
Frago, S., Velázquez-Campoy, A. & Medina, M. (2009). *J. Biol. Chem.* **284**, 6610–6619.
Giancaspero, T. A., Colella, M., Brizio, C., Difonzo, G., Fiorino, G. M., Leone, P., Brandsch, R., Bonomi, F., Iametti, S. & Barile, M. (2015). *Front. Chem.* **3**, 30.

Giancaspero, T. A., Locato, V., de Pinto, M. C., De Gara, L. & Barile, M. (2009). *FEBS J.* **276**, 219–231.
Herguedas, B., Martínez-Júlvez, M., Frago, S., Medina, M. & Hermoso, J. A. (2010). *J. Mol. Biol.* **400**, 218–230.
Huerta, C., Borek, D., Machius, M., Grishin, N. V. & Zhang, H. (2009). *J. Mol. Biol.* **389**, 388–400.
Humphrey, W., Dalke, A. & Schulten, K. (1996). *J. Mol. Graph.* **14**, 33–38.
Kabsch, W. & Sander, C. (1983*a*). *Biopolymers*, **22**, 2577–2637.
Kabsch, W. & Sander, C. (1983*b*). *FEBS Lett.* **155**, 179–182.
Karthikeyan, S., Zhou, Q., Mseeh, F., Grishin, N. V., Osterman, A. L. & Zhang, H. (2003). *Structure*, **11**, 265–273.
Karthikeyan, S., Zhou, Q., Osterman, A. L. & Zhang, H. (2003). *Biochemistry*, **42**, 12532–12538.
Krupa, A., Sandhya, K., Srinivasan, N. & Jonnalagadda, S. (2003). *Trends Biochem. Sci.* **28**, 9–12.
Leskovac, V. (2003). *Comprehensive Enzyme Kinetics.* New York: Kluwer Academic/Plenum Publishers.
Leulliot, N., Blondeau, K., Keller, J., Ulryck, N., Quevillon-Cheruel, S. & van Tilbeurgh, H. (2010). *J. Mol. Biol.* **398**, 641–646.
Marcuello, C., Arilla-Luna, S., Medina, M. & Lostao, A. (2013). *Biochim. Biophys. Acta*, **1834**, 665–676.
Mashhadi, Z., Xu, H., Grochowski, L. L. & White, R. H. (2010). *Biochemistry*, **49**, 8748–8755.
Mashhadi, Z., Zhang, H., Xu, H. & White, R. H. (2008). *J. Bacteriol.* **190**, 2615–2618.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Medina, M. (2012). *Handbook of Flavoproteins*, edited by R. Hille, S. Miller & B. Palfey, pp. 71–100. Berlin: Walter de Gruyter.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Navaza, J. (2001). *Acta Cryst.* D**57**, 1367–1372.
Padilla, J. E. & Yeates, T. O. (2003). *Acta Cryst.* D**59**, 1124–1130.
Park, K.-J., Lee, C.-H., Kim, A., Jeong, K. J., Kim, C.-H. & Kim, Y.-S. (2012). *J. Biol. Chem.* **287**, 3313–3325.
Sandoval, F. J. & Roje, S. (2005). *J. Biol. Chem.* **280**, 38337–38345.
Sandoval, F. J., Zhang, Y. & Roje, S. (2008). *J. Biol. Chem.* **283**, 30890–30900.
Schlitter, J., Engels, M. & Krüger, P. (1994). *J. Mol. Graph.* **12**, 84–89.
Serrano, A., Ferreira, P., Martínez-Júlvez, M. & Medina, M. (2013). *Curr. Pharm. Des.* **19**, 2637–2648.
Serrano, A., Frago, S., Herguedas, B., Martínez-Júlvez, M., Velázquez-Campoy, A. & Medina, M. (2013). *Cell Biochem. Biophys.* **65**, 57–68.
Serrano, A., Frago, S., Velázquez-Campoy, A. & Medina, M. (2012). *Int. J. Mol. Sci.* **13**, 14492–14517.
Serrano, A., Sebastián, M., Arilla-Luna, S., Baquedano, S., Pallarés, M. C., Lostao, A., Herguedas, B., Velázquez-Campoy, A., Martínez-Júlvez, M. & Medina, M. (2015). *Biochim. Biophys. Acta*, **1854**, 897–906.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Walker, R. C., de Souza, M. M., Mercer, I. P., Gould, I. R. & Klug, D. R. (2002). *J. Phys. Chem. B*, **106**, 11658–11665.
Wang, W., Kim, R., Yokota, H. & Kim, S.-H. (2005). *Proteins*, **58**, 246–248.
Yamada, Y., Merrill, A. H. Jr & McCormick, D. B. (1990). *Arch. Biochem. Biophys.* **278**, 125–130.
Yazdanpanah, B., Wiegmann, K., Tchikov, V., Krut, O., Pongratz, C., Schramm, M., Kleinridders, A., Wunderlich, T., Kashkar, H., Utermöhlen, O., Brüning, J. C., Schütze, S. & Krönke, M. (2009). *Nature (London)*, **460**, 1159–1163.
Yruela, I., Arilla-Luna, S., Medina, M. & Contreras-Moreira, B. (2010). *BMC Evol. Biol.* **10**, 311.