The covalent complex of Jo-In results from a long-lived, non-covalent intermediate state with near-native structure

Neil Cox, Cyril Charlier, Ramadoss Vijayaraj, Marion de La Mare, Sophie Barbe, Isabelle André, Guy Lippens, Cédric Montanier

The covalent complex of Jo-In results from a long-lived, non-covalent intermediate state with near-native structure

Neil Cox[1], Cyril Charlier[1], Ramadoss Vijayaraj[2], Marion De La Mare[2], Sophie Barbe[1], Isabelle André[1], Guy Lippens[1*] and Cédric Y. Montanier[1*]

[1] Toulouse Biotechnology Institute (TBI), Université de Toulouse, CNRS, INRAE, INSA, 31077 Toulouse, France

[2] Toulouse White Biotechnology, UMS INRA 1337, UMS CNRS 3582, Institut National des Sciences Appliquées de Toulouse, 31077 Toulouse, France

* Correspondence:

glippens@insa-toulouse.fr, Tel. +33 (0)5 61 55 94 58

and cedric.montanier@insa-toulouse.fr, Tel. +33 (0)5 61 55 97 13.

Abstract

Covalent protein complexes have been used to assemble enzymes in large scaffolds for biotechnology purposes. Although the catalytic mechanism of the covalent linking of such proteins is well known, the recognition and overall structural mechanisms driving the association are far less understood but could help further functional engineering of these complexes. Here, we study the Jo-In complex by NMR spectroscopy and molecular modelling. We characterize a transient non-covalent complex, with structural elements close to those in the final covalent complex. Using site specific mutagenesis, we further show that this non-covalent association is essential for the covalent complex to form.

Keywords

Biomolecular Welding Tool, Jo-In, NMR, Molecular modelling

Introduction

Covalent protein complexes are increasingly used as a tool for building self-assembling multi-enzyme systems [1]. Two commonly used systems are derived from gram-positive bacteria pili, where covalent stabilisation of proteins occurs naturally [2–4]. The first SpyTag/Spycatcher system consists of a minimal β-strand peptide (SpyTag), and a larger

protein fragment (SpyCatcher) derived from the splitting of immunoglobulin-like collagen adhesion domain (CnaB2) from *S. pyogenes* [5]. This complex has been used intensely in biotechnology for building *ex-vivo* enzymatic sub-complexes and scaffolds [6–8]. More recently, a second Jo-In system was presented, based on the RrgA adhesin of *S. pneumoniae.* It consists of two proteins, Jo and In, of similar size (10.5 and 16 kDa, respectively) [9]. This complex has also found some applications in biotechnology, as it allows a simple control on the overall distance between the fused enzymes [10,11].

Both complexes share a similar mechanism for the chemical reaction driving the covalent association. First, a nucleophilic attack by the amine group from a Lysine residue (belonging to SpyTag or Jo) on an Aspartate (for Spycatcher) or Asparagine (for In) residue, is followed by a proton shuttle using the adjacent Glycine or Glutamate residue, and ends with the release of either water or ammonia molecule [12,13]. For this reaction to occur spontaneously, the pK_a of the involved Lys residue needs to be lowered, and this is achieved by formation of a hydrophobic core upon co-folding of the two subunits [14].

Although the catalytic association mechanism for both systems is well characterised, the mechanism driving the protein-protein interaction during complex formation is far less understood. Nuclear Magnetic Resonance (NMR) studies of the SpyTag/Spycatcher complex suggest that their association implies a structural rearrangement of the peptide preceding covalent complex formation, resulting in a transient intermediate that in a second step is further stabilized by the isopeptide bond formation [13,15]. However, whether this mechanism is specific to this system or can be extended to the Jo-In system is still unknown.

Here, we plan to fill this gap in our knowledge of the Jo-In association. By combining molecular modelling and NMR spectroscopy of the Jo protein, free or in complex with its partner In, we aim to better understand the molecular determinants driving the formation of the isopeptide bond. A mutant inactivating the covalent bond is used to decompose the recognition step from the covalent bond formation. The results with this mutant highlight the protein-protein interaction step and structural rearrangement towards the final complex that occurs before covalent bond formation. We further show that the stabilisation of the transient complex implies specific amino acid residues within the interface. Using molecular dynamics simulations and site directed mutagenesis, we show that the formation of this transient complex is essential for covalent complex formation.

Material and methods

Computational procedures

The initial Jo-In complex structure used in the molecular dynamics (MD) study was obtained from the D2 domain of RrgA crystal structure (PDB id: 2WW8) [12] shown to be identical to the structure of the isolated complex (PDB id: 5MKC [9]). Energy minimizations and MD simulations were carried out using AMBER12 suite of programs [16] with the ff12SB and ff99SB force fields [17,18]. The molecular system was neutralized with 7 Na+ ions and solvated with explicit TIP3P water molecules [19], using a cubic box with a minimum distance of 12 Å between the solute and the simulation box edge. Before starting the production simulation, the solvent molecules were energy minimized for 1,000 cycles by restraining the solute atoms followed by 10,000 steps of energy minimization of the whole system without any restraints. The solvent molecules were then equilibrated for 50 ps in the NVT ensemble by restraining the solute atoms. The molecular system was subsequently gradually heated from 100 to 294 K followed by 50 ps of equilibration in the NPT ensemble. During MD preparation procedure, harmonic positional restraints were gradually removed. The final production MD simulation was performed for 10 ns in the NPT ensemble using a 2 fs time step. The temperature (294 K) and the pressure (1 bar) were controlled using the Berendsen algorithms [20]. Particle-Mesh Ewald (PME) summation method [21] was used for calculating long-range electrostatic interactions. The non-bonded interactions were truncated after a 12 Å cut-off. The SHAKE algorithm was used to constrain the lengths of all bonds involving hydrogen atoms to their equilibrium values. Atomic coordinates were saved every 0.4 ps for trajectory analysis using the “cpptraj” program [22] of Ambertools.

All calculations were carried out using Computing MesoCenter of Region Midi-Pyrénées (CALMIP, Toulouse, France). Graphics were prepared using Pymol Molecular Graphics System, (Schrodinger Inc., LLC).

Gene cloning and mutagenesis

Plasmid coding for Jo and In and In^{N695A} were the same as in [9]. The mutant In^{R711E} was prepared from the plasmid coding In protein using site directed mutagenesis (quick change II kit, agilent) Using Primers: 5’-AGAACACAGTG**GAA**GACTTCCCGATT-3’/5’-AATCGGGAAGT**CTT**CCACTGTGTTCT-3’.

Protein expression and purification

All proteins were produced and purified using the same protocol as in [9]. Unlabelled proteins were expressed in BL21(DE3) *Escherichia coli* strain grown to an optical density $O.D._{600nm}$ of 0.6 in LB broth at 37 °C, and expression was induced with 1 mM IPTG for 4 h at 37 °C. Labelled proteins were expressed in BL21(DE3) *Escherichia coli* strain grown to an optical density $O.D._{600nm}$ of 0.6 in LB broth at 37 °C, harvested by centrifugation (4,000 × g, 7 min, 20 °C), transferred to M9 minimal media (40 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.5 mM NaCl, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$) supplemented with 1X MEM vitamins (sigma M6895), appropriate antibiotic and either 20 mM $^{15}NH_4Cl$, 30 mM U-^{13}C Glucose or both and set to grow for an additional 30 min at 37 °C prior induction. In all cases, cells were harvested by centrifugation (5,000 × g, 10 min, 4 °C), resuspended in 20 ml lysis buffer (50 mM sodium phosphate pH 7, 300 mM NaCl, 20 mM imidazole) and stored at -20 °C until purification. Resuspended cell pellets were thawed and incubated 30 min at 30 °C with 0.5 $mg.ml^{-1}$ Lysozyme (euromedex 5934-C) and 1X protease inhibitor cocktail (sigma S8830). Cells were lysed by sonication (amplitude 25, 1 min on/ 2 min rest, repeated twice) and lysate was clarified by centrifugation (60,000 × g, 30 min, 4 °C). Expressed polyhistidine-tagged proteins were purified by immobilized metal ion affinity chromatography (IMAC) using 3 ml of Talon resin (TALON® Metal Affinity Resin, Clontech) previously equilibrated in lysis buffer. After application of clarified lysate, resin was washed with 10 column volume (CV) of lysis buffer, and the protein of interest was eluted using an increasing concentration of imidazole (80 mM, 100 mM and 300 mM). Purified protein purity was controlled by SDS-PAGE (Any kD™ Mini-PROTEAN® TGX Stain-Free™ Protein Gels, Bio-rad). Samples were subsequently buffer exchanged against 50 mM sodium phosphate buffer, pH 7 using PD-10 column (GE). Protein concentrations were determined by measuring absorbance at 280nm (nanodrop 2000, Thermo scientific). Theoretical molar extinction coefficients and molecular weight were calculated using ProtParam online software (10,601.5 $M^{-1}.cm^{-1}$ for Jo and 16,449.1 $M^{-1}.cm^{-1}$ for In) ([23]). Proteins were concentrated using centrifugal concentrator units (Amicon® Ultra 3K, Merck KGaA). Aliquots were stored at -20 °C until use.

NMR analysis

All NMR experiments were performed at 303°K on a Bruker Avance III HD 800 MHz spectrometer equipped with a 5 mm quadruple resonance QCI-P (H/P-C/N/D) cryogenically

cooled probe head (Bruker Biospin, Billerica, USA). Spectra were acquired and processed using the Bruker Topspin 3.5 software. ^{1}H-^{15}N TROSY spectra (Pervushin et al., 1997) were acquired by recording 2048 direct time points with 16 transients and 128 indirect time points. Spectra were zero filled to 8192 direct and 512 indirect time points and apodized with a $\pi/2$ shifted square sine window function before Fourier transform. Triple resonance experiments (HNCO, HNCACO, HNCA and HNCACB) were recorded on either 350 µM free dually labelled Jo protein or 150 µM of the same protein in a covalent complex with unlabelled In protein. A total of 16 transients were recorded with 2048 time points in the direct acquisition, and 64 (F2) × 64 (F1) indirect time points for HNCO, HNCA and HNCACO and 64 × 320 indirect time points for HNCACB. All 3D experiments were transformed as a matrix with 2048 × 512 ×512 timepoints.

Results and discussion

The Jo/In system features two protein fragments (composed of 79 and 133 residues, respectively) capable of spontaneously and irreversibly crosslink through an isopeptide bond between side chains of K_{191} of Jo and N_{695} of In and involving Asp_{600} of In as a proton shuttle [9]. Figure 1 shows the ^{1}H, ^{15}N TROSY spectrum of the ^{15}N labelled Jo protein, free or covalently bound to unlabelled In. The first spectrum contains peaks mostly distributed within a narrow ^{1}H chemical shift range between 8-8.5 ppm. The failure of isolated Jo protein to crystallise (Bonnet, personal communication) and the reduced chemical shift range of its amide protons all suggest a lack of stable tertiary structure and pronounced flexibility of the protein in its free state [24]. The spectrum of the bound protein, in contrast, shows peaks spread over a larger chemical shift range, suggesting that the Jo protein structure changes upon engaging in its covalent complex with In towards a more globular protein, in agreement with the crystallographic structure of the complex (PDB id: 5MKC) [9].

Using triple resonance experiments on the ^{15}N, ^{13}C labelled Jo protein free or in its covalent complex with unlabelled In, the assignment of Jo in both conditions was attempted. Despite its small size, full assignment of the spectra of Jo when free in solution proved impossible, due to extensive line broadening of several signals in the 3D experiments. We obtained with confidence the assignment of about half of the protein, comprising the C-terminal β-sheet (composing the interface with In in the covalent complex) and several loops throughout the protein. Extensive NMR line broadening has been described for molten

globule proteins [25,26] and suggests that the isolated Jo would already contain some structural elements. Those amino acids that will later constitute the interface of the complex, however, are devoid of structure and give therefore rise to sharp spectra. Unexpectedly, extensive line broadening was equally observed for many resonances in the spectrum of the covalent Jo-In complex, allowing only the full assignment of its two C-terminal β-sheets, composing the interface with In (Figure 2.c). Therefore, even in the final covalent complex, significant flexibility is maintained in parts of the protein. Important for the recognition mechanism, our NMR data show that the Jo C-terminal strand evolves from flexible in the free protein to a rigid β-sheet in the complex.

Covalent bond formation is not essential for Jo-In interaction.

Data from previous studies suggest that covalent complex formation is preceded by at least partial formation of a hydrophobic pocket that shields the future isopeptide bond from water and lowers the pKa of the catalytic Lysine residue [13,14]. To further investigate this mechanism, we studied the behaviour of the labelled Jo protein upon addition of an inactive mutant partner In^{N695A}. As in this protein N_{695}, involved in the covalent linkage, is mutated [9], we used it to monitor whether a reversible intermediate complex can form without evolving towards the final covalent complex. In the ^{1}H, ^{15}N TROSY spectrum of labelled Jo after addition of one molar equivalent of In^{N695A}, we observe specific peaks (green peaks in Figure 2.a) at chemical shift values exactly or close to that of Jo peaks in the spectrum of the covalent complex (in grey). Our partial assignment indicated that those peaks correspond all to residues on the two C-terminal β-sheets that form the interface with In (T_{188}-S_{190}, Q_{195}-N_{198}, E_{207}, S_{211}-G_{212}, T_{214}-V_{215}, Figure 2.c). In the covalent Jo-In complex, all these residues physically interact with amino acid residues from In, suggesting that they are key players of the interface. Their proximity in the non-covalent Jo-In complex to the peaks of Jo in its covalent complex suggests that this transient structure resembles closely that of the covalent complex.

We further observe peaks corresponding to the resonances of the free protein (in blue). Their significant intensity points to an important fraction of free Jo not interacting with In^{N695A}. The transient complex being the first step of Jo-In covalent association hence would result from a low affinity interaction. Yet, for such a low affinity interaction, exchange between bound and free forms is generally fast, so we would expect for a given residue a cross peak at a population-averaged chemical shift value rather than two distinct peaks.

Moreover, peaks corresponding to the non-covalent complex increase in intensity upon titration of In^{N695A} up to 5 equivalents of Jo, but do not move (Figure 2.b). The non-covalent complex can hence be described as a low-affinity complex characterized by a slow exchange rate between bound and free forms. The small proportion of this intermediate complex further explains the long incubation time needed for Jo-In covalent association (in the range of hours [6]); the correctly folded non-covalent complex is essential for the covalent chemistry to occur, but is a rare event that therefore takes time. Diffusion controlled association of the non-covalent complex in contrast is expected to lead to a total complex formation within seconds.

Identification of amino acids engaged in recognition and complex stabilisation

Molecular dynamics (MD) simulations were carried out to further understand the molecular determinants controlling association of Jo and In, as well as the stability of the system.

Analysis of the Jo/In interface revealed that all amino acid residues from the Jo protein are accessible from the surface and 68 % of its residues are engaged in the interface with In. The interface is not symmetrical, as In provides only 36 % of its residues to the interface. Altogether, nearly 100 amino acid residues from both Jo and In participate to the interface.

MD simulation (10 ns) performed on the covalent Jo/In complex further confirmed the important stability of the complex, with limited root mean square deviation (rmsd) of the backbone and side chain heavy atoms of Jo-In complex with respect to the initial structure (Figure 3.a) and absence of any major conformational rearrangement during the course of the simulation (Figure 3.b). Detailed analysis of the H-bonding interactions along the MD trajectory revealed that on average, Jo and In proteins can form 26 intermolecular H-bonding interactions (Figure 3.c). In addition to the intermolecular bonding interactions, a water molecule at the interface of Jo and In forms a hydrogen bond with the isopeptide bond. Finally, E_{186} (Jo)/R_{711} (In) and D_{201} (Jo)/R_{681} (In) further form tight ionic interactions that stabilize the complex during the MD trajectory.

Transient complex formation is required for covalent bond formation

The E_{186} (Jo)/R_{711} (In) ionic interaction is close to the part of the β-strand that according to the NMR analysis of the non-covalent complex adopts a near-native conformation even in

the absence of the isopeptide bond. This orthogonal pair hence might be one of the driving factors initiating and/or stabilizing the non-covalent interaction between Jo and In that precedes covalent complex formation. Using site directed mutagenesis, R_{711} of In was mutated to E, which should disrupt this electrostatic bond. However, N_{695} was maintained in In, thereby maintaining its intrinsic catalytic ability to form a covalent complex. Upon mixing ^{15}N labelled Jo with this In^{R711E}, its $^{1}H,^{15}N$ spectrum perfectly overlaps with that of isolated Jo, and does not show additional peaks corresponding to a complex, be it covalent or not (Figure 4.a). Furthermore, SDS-PAGE analysis shows the failure of both proteins to form the covalent complex on the same two-hour time scale where the wild type pair forms (Figure 4.b). Only after overnight incubation of both proteins, a slight band could be detected on the SDS-PAGE gel, at the expected size of the covalent Jo-In complex (Figure 4.b).

Breaking the ionic interaction hence does not impair covalent bond formation through interfering with the chemistry of bond formation, but rather renders the statistical event of forming a non-covalent pair that can evolve into a covalent complex even rarer than for the wild type pair. This charged pair thus plays a central role in the first recognition steps of the complex that ultimately leads to covalent bond formation.

Conclusion

Our present study has elucidated some key aspects of the Jo-In covalent complex formation. First, Jo starts as a protein devoid of a stable secondary or tertiary structure, notably for its C-terminal strand that will form the interface with In in the covalent complex (Figure 1). Upon interaction with In, it can form a well-structured covalent complex as characterized by X-ray crystallography, but mobility is still a defining factor as indicated by the extensive line broadening we observe (Figure 1). When interfering with the covalent isopeptide bond formation through mutation of the catalytic N_{695} of In, Jo cross peaks in this non-covalent Jo-In^{N695A} complex coinciding with cross peaks of corresponding residues in the Jo-In covalent complex suggest that a non-covalent complex with a properly folded In interface forms before and independently from the isopeptidic bond (Figure 2). This non-covalent complex is rare, but sufficiently long-lived to be observed as distinct cross peaks in the ^{1}H, ^{15}N spectrum rather as a population averaged spectrum (Figure 2). Closer investigation by molecular modelling allowed identification of the E_{186}(Jo)/R_{711}(In) pair as a potential driver for this non-covalent complex formation. Mutation of the latter R_{711} residue of In proved

indeed deleterious with the first step of non-covalent association, comforting the hypothesis that a long lived partially structured intermediate complex is required for covalent bond formation.

Our study thus suggests that the intermediate, non-covalent complex structure of the Jo-In pair is close to its final covalent complex structure, which was not the case for the SpyTag-Spycatcher complex [15]. This difference could be explained by the small size of the SpyTag peptide, whereby the interface is minimal, whereas it is much larger for the Jo-In complex.

Acknowledgements

We thank Thierry Vernet and Anne-Marie Di Guilmi (Institut de Biologie Structurale, Grenoble, France) for providing the plasmids coding for Jo and In. We thank Dr E Cahoreau and L. Peyriga for expert support of the NMR facility. MetaToul (Toulouse metabolomics & fluxomics facilities, www.metatoul.fr) is part of the French National Infrastructure for Metabolomics and Fluxomics MetaboHUB-AR-11-INBS-0010 (www.metabohub.fr), and is supported by the Région Midi-Pyrénées, the ERDF, the SICOVAL and the French Minister of Education & Research, who are all gratefully acknowledged.

Funding

This work was partially funded by an internal grant from the Toulouse White Biotechnology (mIMHETIQ project, 2014-2016) and was granted access to the HPC resources on the Computing mesocenter of Région Midi-Pyrénées (CALMIP, Toulouse, France).

References

[1] S.-Z. Wang, Y.-H. Zhang, H. Ren, Y.-L. Wang, W. Jiang, B.-S. Fang, Strategies and perspectives of assembling multi-enzyme systems, Crit. Rev. Biotechnol. 37 (2017) 1024–1037. https://doi.org/10.1080/07388551.2017.1303803.

[2] T. Proft, E.N. Baker, Pili in Gram-negative and Gram-positive bacteria — structure, assembly and their role in disease, Cell. Mol. Life Sci. 66 (2009) 613–635. https://doi.org/10.1007/s00018-008-8477-4.

[3] H.J. Kang, E.N. Baker, Structure and assembly of Gram-positive bacterial pili: unique covalent polymers, Curr. Opin. Struct. Biol. 22 (2012) 200–207. https://doi.org/10.1016/j.sbi.2012.01.009.

[4] J.L. Telford, M.A. Barocchi, I. Margarit, R. Rappuoli, G. Grandi, Pili in Gram-positive pathogens, Nat. Rev. Microbiol. 4 (2006) 509–519. https://doi.org/10.1038/nrmicro1443.

[5] B. Zakeri, J.O. Fierer, E. Celik, E.C. Chittock, U. Schwarz-Linek, V.T. Moy, M. Howarth, Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin, Proc. Natl. Acad. Sci. 109 (2012) E690–E697.

[6] Z. Liu, S. Cao, M. Liu, W. Kang, J. Xia, Self-Assembled Multienzyme Nanostructures on Synthetic Protein Scaffolds, ACS Nano. 13 (2019) 11343–11352. https://doi.org/10.1021/acsnano.9b04554.

[7] Z. Botyanszki, P.K.R. Tay, P.Q. Nguyen, M.G. Nussbaumer, N.S. Joshi, Engineered catalytic biofilms: Site-specific enzyme immobilization onto *E. coli* curli nanofibers: Catalytic Biofilms Using Engineered Curli Fibers, Biotechnol. Bioeng. 112 (2015) 2016–2024. https://doi.org/10.1002/bit.25638.

[8] G. Zhang, M.B. Quin, C. Schmidt-Dannert, Self-Assembling Protein Scaffold System for Easy in Vitro Coimmobilization of Biocatalytic Cascade Enzymes, ACS Catal. 8 (2018) 5611–5620. https://doi.org/10.1021/acscatal.8b00986.

[9] J. Bonnet, J. Cartannaz, G. Tourcier, C. Contreras-Martel, J.P. Kleman, C. Morlot, T. Vernet, A.M. Di Guilmi, Autocatalytic association of proteins by covalent bond formation: a Bio Molecular Welding toolbox derived from a bacterial adhesin, Sci. Rep. 7 (2017). https://doi.org/10.1038/srep43564.

[10] C.Y. Montanier, M. Fanuel, H. Rogniaux, D. Ropartz, A.-M. Di Guilmi, A. Bouchoux, Changing surface grafting density has an effect on the activity of immobilized xylanase towards natural polysaccharides, Sci. Rep. 9 (2019). https://doi.org/10.1038/s41598-019-42206-w.

[11] T. Enjalbert, M. De La Mare, P. Roblin, L. Badruna, T. Vernet, C. Dumon, C.Y. Montanier, Characterisation of the Effect of the Spatial Organisation of Hemicellulases on the Hydrolysis of Plant Biomass Polymer, Int. J. Mol. Sci. 21 (2020) 4360. https://doi.org/10.3390/ijms21124360.

[12] T. Izoré, C. Contreras-Martel, L. El Mortaji, C. Manzano, R. Terrasse, T. Vernet, A.M. Di Guilmi, A. Dessen, Structural Basis of Host Cell Recognition by the Pilus Adhesin from Streptococcus pneumoniae, Structure. 18 (2010) 106–115. https://doi.org/10.1016/j.str.2009.10.019.

[13] R.M. Hagan, R. Björnsson, S.A. McMahon, B. Schomburg, V. Braithwaite, M. Bühl, J.H. Naismith, U. Schwarz-Linek, NMR Spectroscopic and Theoretical Analysis of a Spontaneously Formed Lys-Asp Isopeptide Bond, Angew. Chem. Int. Ed. 49 (2010) 8421–8425. https://doi.org/10.1002/anie.201004340.

[14] X. Hu, H. Hu, J.A. Melvin, K.W. Clancy, D.G. McCafferty, W. Yang, Autocatalytic Intramolecular Isopeptide Bond Formation in Gram-Positive Bacterial Pili: A QM/MM Simulation, J. Am. Chem. Soc. 133 (2011) 478–485. https://doi.org/10.1021/ja107513t.

[15] N. Zhang, J. Liu, Y. Liu, W.-H. Wu, J. Fang, X. Da, S. wang, W.-B. Zhang, NMR Spectroscopic Studies Reveal the Critical Role of Isopeptide Bond in Forming the Otherwise Unstable SpyTag-SpyCatcher Mutant Complexes, Biochemistry. 59 (2020) 2226–2236. https://doi.org/10.1021/acs.biochem.0c00287.

[16] D.A. Case, T.A. Darden, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Götz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe,

D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman, AMBER 12, Univ. Calif. San Franc. (2012).

[17] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, Comparison of multiple Amber force fields and development of improved protein backbone parameters, Proteins Struct. Funct. Bioinforma. 65 (2006) 712–725. https://doi.org/10.1002/prot.21123.

[18] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB, J. Chem. Theory Comput. 11 (2015) 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255.

[19] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935. https://doi.org/10.1063/1.445869.

[20] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690. https://doi.org/10.1063/1.448118.

[21] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An *N* ·log(*N*) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092. https://doi.org/10.1063/1.464397.

[22] D.R. Roe, T.E. Cheatham, PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data, J. Chem. Theory Comput. 9 (2013) 3084–3095. https://doi.org/10.1021/ct400341p.

[23] M.R. Wilkins, E. Gasteiger, A. Bairoch, J.-C. Sanchez, K.L. Williams, R.D. Appel, D.F. Hochstrasser, Protein Identification and Analysis Tools in the ExPASy Server, in: 2-Proteome Anal. Protoc., Humana Press, New Jersey, 1998: pp. 531–552. https://doi.org/10.1385/1-59259-584-7:531.

[24] D.S. Wishart, B.D. Sykes, The 13 C chemical-shift index: a simple method for the identification of protein secondary structure using 13 C chemical-shift data, J. Biomol. NMR. 4 (1994) 171–180.

[25] D. Eliezer, P.E. Wright, Is apomyoglobin a molten globule? Structural characterization by NMR, J. Mol. Biol. 263 (1996) 531–538. https://doi.org/10.1006/jmbi.1996.0596.

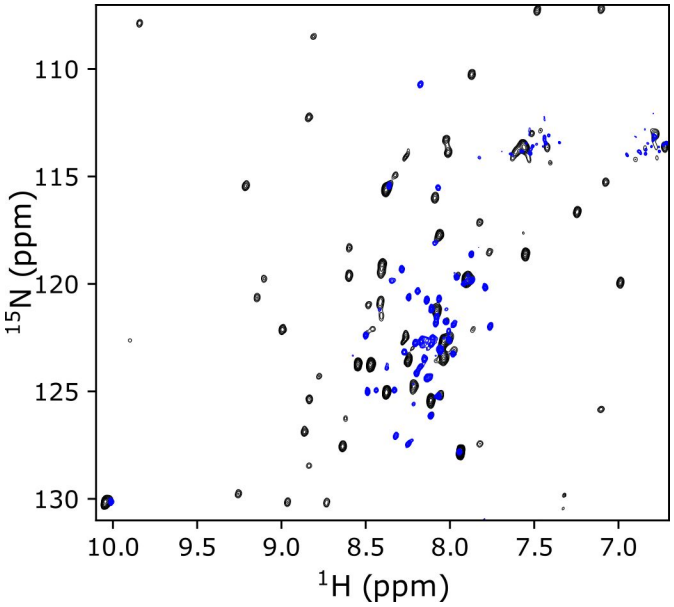
[26] S. Cavagnero, C. Nishimura, S. Schwarzinger, H.J. Dyson, P.E. Wright, Conformational and dynamic characterization of the molten globule state of an apomyoglobin mutant with an altered folding pathway, Biochemistry. 40 (2001) 14459–14467. https://doi.org/10.1021/bi011500n.

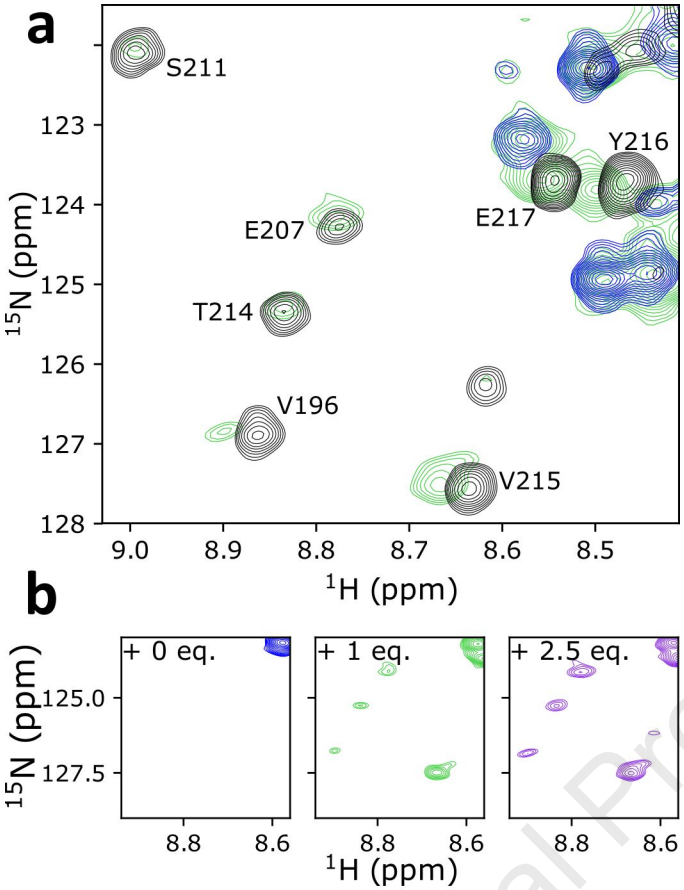
Figure 1.a. Overlay of the ^{1}H, ^{15}N TROSY spectra measured on 150 µM of the free Jo (blue) or in its covalent complex after incubation with an excess (180 µM) of unlabelled In protein (black).

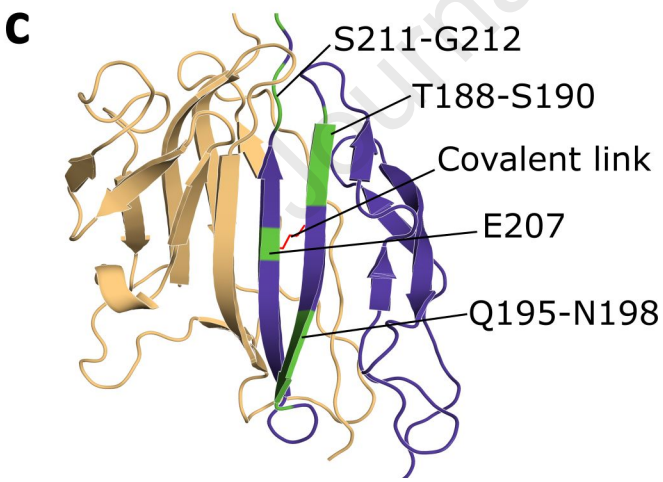
Figure 2. a. Overlay of the ^{1}H, ^{15}N TROSY spectra of the labelled Jo protein (100 µM), free (blue), in interaction with one equivalent of its inactive mutant partner In^{N695A} (green), or after formation of its covalent complex with In (black). Indicated residues are assigned in the latter spectrum. **b.** Titration of labelled Jo with In^{N695A}. Blue: no In^{N695A}. Green: one equivalent. Purple: 5 equivalents of In^{N695A}. **c.** Structure of the Jo-In complex (blue and orange, respectively) from pdb (5MKC). Residues for which the signal appear upon transient interaction and could be assigned are shown in green (indicated residues in panel a).

Figure 3. a. Plot of backbone (black) and sidechain (red) rmsd of Jo-In complex with respect to the initial structure during the 10 ns MD simulation. **b**. Superimposed initial and final structures of Jo-In complex from the MD simulation. The isopeptide bond is shown in yellow color. **c.** Pairs of amino acid residues that form intermolecular H-bonding interaction as observed in the Jo-In initial structure.

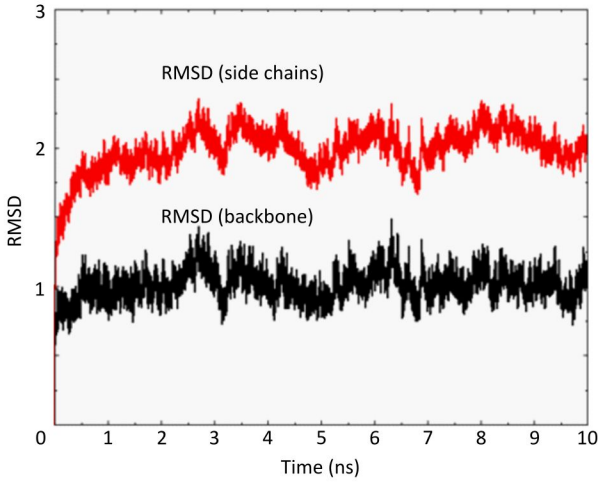
Figure 4.a. Overlay of the ^{1}H, ^{15}N TROSY spectra of the Jo protein (150 µM) in its free state (dark blue) or upon addition of one equivalent In^{R711E} (light blue). **b.** SDS-PAGE analysis after one hour following mixing Jo with In or In^{R711E}. Lanes: After 2h incubation (1-6): 1: Standard (kDa); 2: Jo; 3: In; 4: In^{R711A}; 5: Jo+In; 6: Jo+In^{R711E}. After overnight incubation (8-12): 8: Jo. 9: In; 10: In^{R711A}; 11: Jo+In; 12: Jo+In^{R711E}; Expected size for covalent complex bands are indicated between dashes.



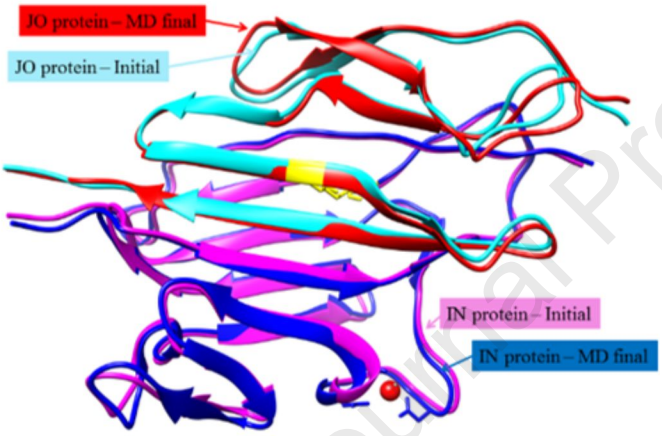




a

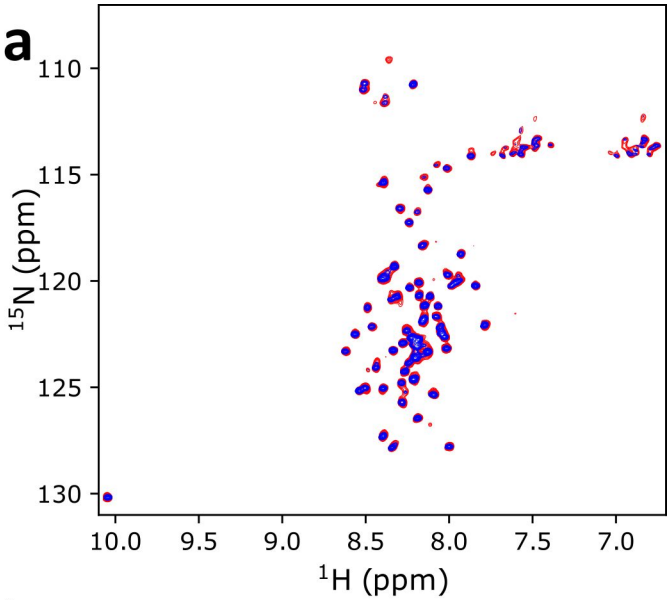


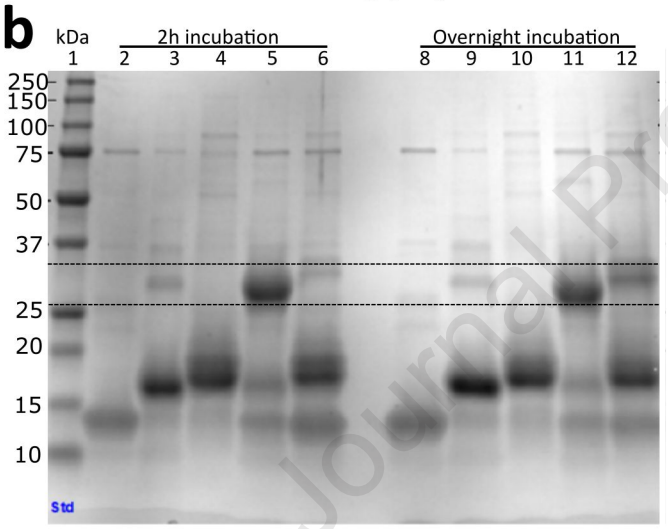
b



c

Nr.	JO	IN
Backbone-Backbone		
1	L 208	L 676
2	I 206	Y 678
z		
3	L 175	K 717
4	V 210	V 674
5	L 213	S 592
6	V 215	K 590
7	Y 204	V 680
8	G 212	E 672
9	N 202	L 682
10	K 173	R 719
Sidechain-Sidechain		
11	S 142	R 719
12	N 197	N 679
13	K 173	D 720
14	N202	D 720
15	G 186	R 711
16	D 201	R 681
Backbone-Sidechain Sidechain-backbone		
17	T 214	T 670
18	S 142	N 689
19	S 211	E 672
20	Y 158	I 715
21	S 142	R 719
22	K 213	S 592
23	E 186	S 592
24	D 153	D 712
25	R 197	D 712
26	H 172	R 719





Highlights

- The covalent bond formation is not essential for Jo-In interaction
- A specific couple of amino acids at the Jo/In interface is engaged in the recognition and the stabilization of the complex
- A transient complex formation is required for covalent bond formation between Jo and In

The covalent complex of Jo-In results from a long-lived, non-covalent intermediate state with near-native structure

The authors (Neil Cox, Cyril Charlier, Ramadoss Vijayaraj, Marion De La Mare, Sophie Barbe, Isabelle André, Guy Lippens and Cédric Y. Montanier) declare that there is no conflict of interest