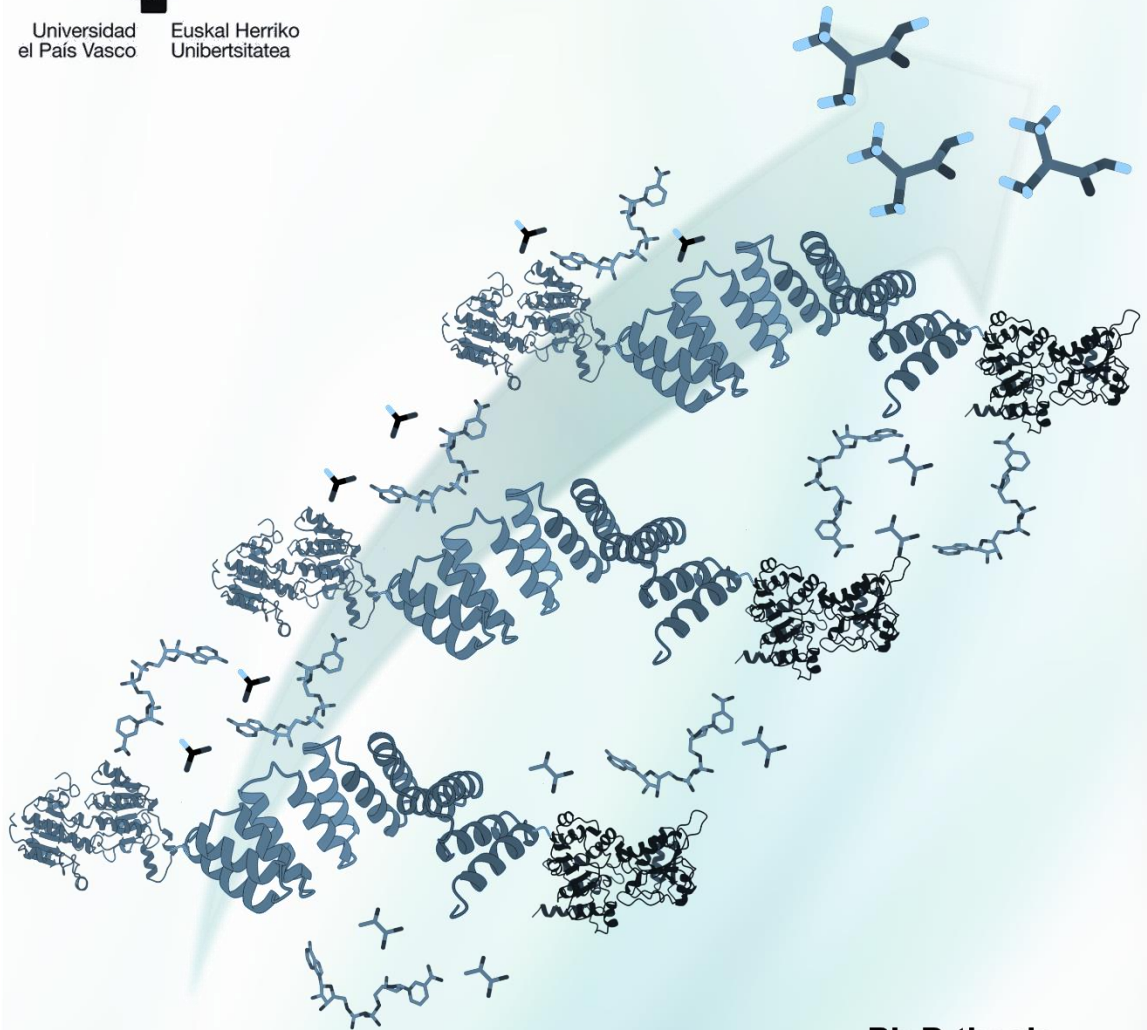


eman ta zabal zazu



Universidad
el País Vasco

Euskal Herriko
Unibertsitatea



Ph.D thesis

Novel organized multi-enzyme systems based on protein scaffolding

Alba Ledesma Fernandez

Donostia, 2023



Novel organized multi-enzyme systems based on protein scaffolding

Ph.D Thesis

To obtain the Doctor Philosophy degree in Applied Chemistry and Polymers at the University of the Basque Country (EHU/UPV)

By

Alba Ledesma Fernandez

Donostia-San Sebastián, 2023

Thesis supervisors: Dra. Aitziber López Cortajarena (Biomolecular Nanotechnology Laboratory, CIC biomaGUNE)

Dr. Fernando López Gallego (Heterogeneous biocatalysis Laboratory, CIC biomaGUNE)

University tutor: Dra. Zoraida Freixa Fernandez (Department of Applied Chemistry, University of the Basque Country)

Novel organized multi-enzyme systems based on protein scaffolding

Ph.D Thesis

To obtain the Doctor Philosophy degree in Applied Chemistry and Polymers at the University of the Basque Country (EHU/UPV)

By

Alba Ledesma Fernandez

Donostia-San Sebastián, 2023

Thesis supervisors: Dra. Aitziber López Cortajarena (Biomolecular Nanotechnology Laboratory, CIC biomaGUNE)

Dr. Fernando López Gallego (Heterogeneous biocatalysis Laboratory, CIC biomaGUNE)

University tutor: Dra. Zoraida Freixa Fernandez (Department of Applied Chemistry, University of the Basque Country)

Funding Agencies

This thesis was financial support provided by Era-CoBioTech (Project ID: 61 HOMBIOCAT/ PCI2018-092984) and by the Agencia Estatal de Investigación, Spain (PID2019-111649RB-I00). This work was performed under the Maria de Maeztu Units of Excellence Program from the Spanish State Research Agency Grant MDM-2017-0720 (CIC biomaGUNE).

The following results from this thesis have already been published and are given below:

1) **Ledesma-Fernandez, A.**, Velasco-Lozano, S., Santiago-Arcos, J. *et al.* Engineered repeat proteins as scaffolds to assemble multi-enzyme systems for efficient cell-free biosynthesis. *Nat Commun* **14**, 2587 (2023). <https://doi.org/10.1038/s41467-023-38304-z>. Parts of chapters 3 and 5 correspond to this publication.

Table of contents

Resumen	15
Summary	25
Chapter 1. General Introduction	33
1.1. Biocatalysis	35
1.2. Enzyme immobilization on supports	38
1.2.1. Solid supports	39
1.2.2. Biomacromolecule-based supports	41
1.3. Proteins as building blocks for enzyme immobilization: general strategies, spatial organization and substrate channeling effect	42
1.4. Enzyme immobilization onto protein-based scaffolds	45
1.5. Repeat proteins as ideal scaffolds for enzyme immobilization	48
Chapter 2. General aim and objectives	53
Chapter 3. Experimental procedures	59
3.1. Protein design, Cloning and Molecular Biology	61
3.1.1. SCAB _C modules	61
3.1.2. SCAB _H modules	62
3.1.3. SCAB-enzyme fusions	64
3.1.4. TRAP scaffolds and tagged enzymes	66
3.2. Protein expression and purification	69
3.3. MALDI-TOF mass spectrometry	71
3.4. Circular dichroism	71
3.5. SCAB _C modules and SCAB _C -enzyme fusions assembly	72
3.6. SCAB _H modules and SCAB _H -enzyme fusions assembly	73
3.7. Enzyme assembly onto TRAP proteins	73
3.8. Determination of the Hydrodynamic radius (Rh)	74
3.8.1. SCAB _H assembly	75
3.8.2. AlaDH/FDH@TRAP1-3 assembly	75
3.9. Size determination by size exclusion chromatography (SEC)	76
3.10. Inductively coupled plasma mass spectrometry (ICP-MS)	76
3.11. Enzymatic activity measurements	77
3.12. Enzyme kinetic parameters	77
3.13. Biotransformation of L-Alanine	78
3.14. Fluorescence anisotropy-based binding assay	79
3.15. Competitive side reaction catalytic assay	80
3.16. Isotope enrichment and dilution assay: deuterated and non-deuterated L-Alanine product formation	80
3.17. Fluorescence Confocal Microscopy	81
3.18. L-Alanine synthesis with heterogeneous biocatalysts	81
3.19. Benzylamine synthesis catalyzed by scaffolded three-enzyme multi-enzymatic system	82
Chapter 4. Engineering Bio-brick protein scaffolds for enzyme assembly	83
4.1. Introduction	85
4.1.1. Enzyme assembly via biomacromolecules: fusion proteins	85

Table of contents

4.1.2. CTPR proteins as scaffolding units	87
4.1.3. FDH/LAlaDH bi-enzyme cascade on CTPR proteins	88
4.2. Results and discussion	89
4.2.1. Design of the SCAffolding Bricks (SCABs)	89
4.2.2. Assembly, catalytic activity, and biosynthesis of L-Alanine of the multi-enzyme complexes through engineered SCAB Bio-bricks	101
4.3. Conclusions	110
Chapter 5. Engineered repeat proteins as scaffolds for assembling multi-enzyme systems for efficient cell-free biosynthesis	111
5.1. Introduction	113
5.1.1. Cell-free biosynthesis	113
5.1.2. Scaffolded enzymes: spatial arrangement and channeling effect	114
5.1.3. TRAP proteins as scaffolding units	115
5.1.4. FDH/ ω TA/LAlaDH tri-enzyme cascade on TRAP proteins	116
5.2. Results and discussion	117
5.2.1. Design of the scaffolding strategy to assemble multi-enzyme systems	117
5.2.2. Catalytic activity of multi-enzymatic systems	123
5.2.3. Biosynthesis of L-Alanine catalyzed by TRAP-scaffolded enzyme system	127
5.2.4. Characterization of the channeling effect by the scaffolded multi-enzyme systems	132
5.2.5. Heterogenization of the enzyme scaffolds	135
5.2.6. Amine biosynthesis catalyzed by a TRAP-scaffolded tri-enzyme system	142
5.3. Conclusions	145
General conclusions	147
References	153
Abbreviations	165
Acknowledgments	171

Resumen

La biocatálisis es el campo científico relacionado con el uso de enzimas como catalizadores en muchas conversiones químicas importantes en la industria. Las enzimas presentan ventajas extraordinarias en comparación con sus competidores directos, es decir, los catalizadores químicos, como un mayor rendimiento catalítico, selectividad y especificidad. Estas propiedades las convierten en candidatas ideales para catalizar reacciones en la industria que no pueden llevarse a cabo con catalizadores convencionales, como la conversión completa de los sustratos en un solo paso. Sin embargo, el principal inconveniente de las enzimas, además de su elevado coste, es su susceptibilidad a la desnaturalización cuando se exponen a condiciones adversas, como altas temperaturas. Esta desnaturalización reduce su estabilidad operativa, afectando su aplicabilidad y provocando una disminución en su actividad catalítica.

En los últimos años, se han realizado esfuerzos para abordar los retos asociados a la inmovilización de enzimas, explorando las ventajas de inmovilizar enzimas en soportes, centrándose específicamente en la organización espacial de las enzimas. Aunque la inmovilización en soportes sólidos es un método robusto, presenta importantes inconvenientes, como la dificultad de co-inmovilizar numerosas enzimas, la lentitud química requerida, o la falta de control en la organización en escalas submicromolares. Sin embargo, en la última década, han ganado popularidad tecnologías más modernas de soporte de enzimas, como el uso de biomacromoléculas como el ADN, péptidos, nanopartículas y proteínas. Las ventajas de estos soportes son la capacidad de inmovilizar varias enzimas, el control espacial en diferentes escalas de tamaño, el control de la interacción de los sustratos y cofactores con los sistemas soportados, y la mejor comprensión y el control de las reacciones que tienen lugar en la cascada enzimática. Al conseguir una organización espacial específica de las enzimas a nivel nanométrico, los soportes basados en biomacromoléculas, no solo proporcionan control espacial, sino que también mejoran de la eficacia catalítica y permiten efectos de canalización de sustratos y cofactores.

En esta tesis, además de abordar las limitaciones anteriormente descritas de los soportes sólidos, hemos ido un paso más allá utilizando andamiajes de proteínas sintéticas de diseño como soportes de enzimas. Estos andamiajes basados en proteínas de diseño TPR (del inglés *Tetratricopeptide repeat*) han sido ingenierizadas para permitir el ensamblaje ordenado en escala nanométrica de enzimas. Las principales características de estos andamiajes son su modularidad y su estructura simplificada y conocida que permiten su modificación sin comprometer su estructura, así como sus propiedades intrínsecas de autoensamblaje supramolecular que les permiten ensamblarse y proporcionar un control espacial preciso a nanoescala. Estos sistemas por lo tanto superan en ventajas a los andamiajes proteicos convencionales. Por último, estos andamiajes proteicos, debido a la capacidad de ingenierizar sus propiedades, pueden codificar interacciones controladas con cofactores esenciales, sustratos, o intermediarios de reacción, lo que los hace muy prometedores para diversas aplicaciones en el campo de la biocatálisis.

Los dos enfoques seleccionados para organizar las enzimas en el espacio en escala nanométrica son el ensamblaje supramolecular a través de propiedades intrínsecas de autoensamblado en estructuras lineales de los andamiajes seleccionados y el

ensamblaje basado en el reconocimiento biomolecular. La estabilización de los módulos CTPR (del inglés *Consensus Tetratricopeptide repeat*) diseñados denominados como SCAB (del inglés *SCAffolding Bio-Bricks*) y ensamblados mediante interacciones intermoleculares cabeza-cola se consigue implementando enlaces disulfuro mediados por cisteína y ensamblaje dirigido por metales. Por último, para el ensamblaje mediante reconocimiento biomolecular, elegimos los módulos TRAP (del inglés *Tetratricopeptide Repeat Affinity Protein*) con una capacidad de reconocimiento específica y ortogonal frente a sus péptidos diana (Figura R1).

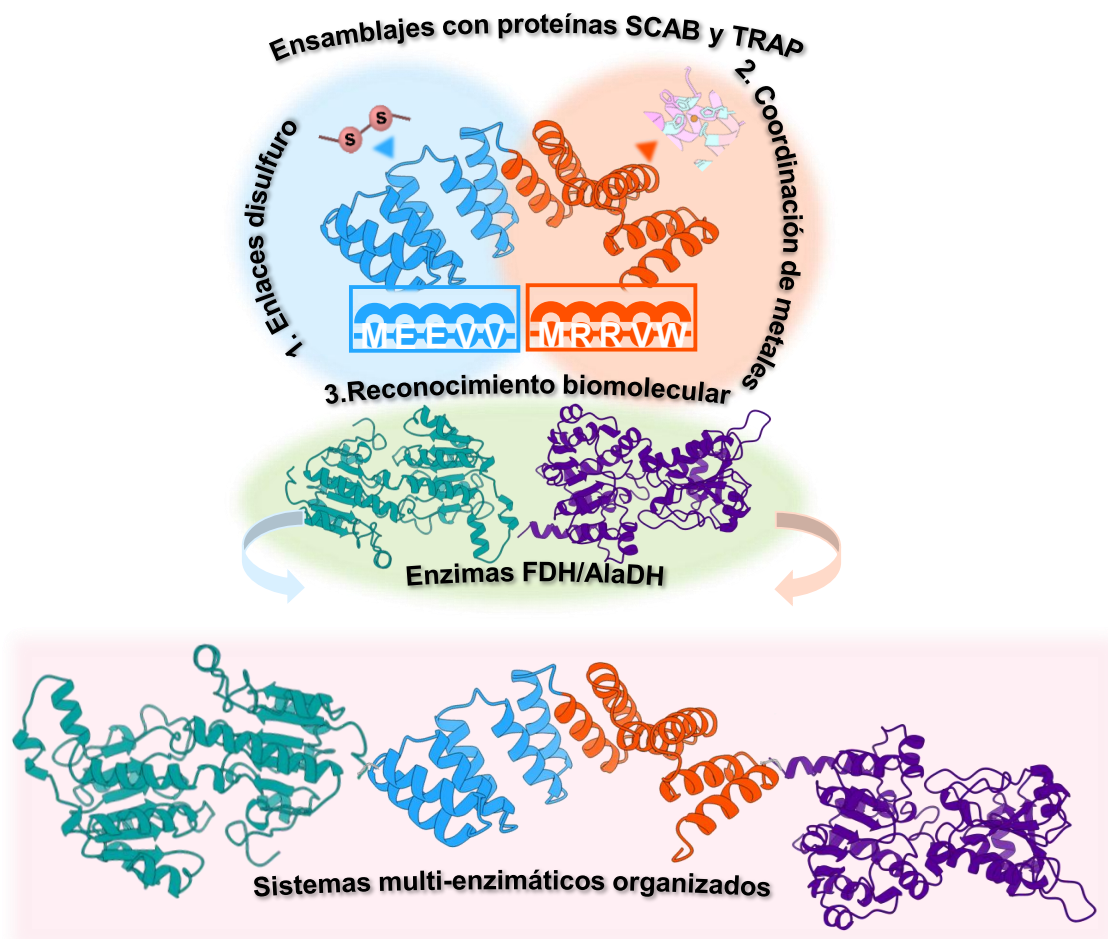


Figura R1. Esquema de los nuevos sistemas multi-enzimáticos organizados basados en las proteínas SCAB y TRAP. Tres estrategias de ensamblaje enzimático: módulos de autoensamblaje SCAB fijados a través de enlaces disulfuro mediados por cisteína (1) o a través de coordinación de metales (2) y módulos TRAP para el ensamblaje a través del reconocimiento biomolecular (3).

En concreto, los andamiajes proteicos utilizados en esta tesis son dos variantes sintéticas originadas a partir del dominio proteico natural de repetición de tetratricopéptido (TPR, del inglés *tetratricopeptide repeat*): dominios de autoensamblaje basados en la secuencia TPR consenso (CTPR) y diseñados para estabilizar los autoensamblados por diferentes vías (SCAB) y dominios ingenierizados con propiedades de bioreconocimiento (TRAP).

La secuencia CTPR consenso está compuesto por 34 aminoácidos de los cuales sólo 8 son altamente conservados en secuencia y definen su plegamiento característico dando lugar a una estructura de hélice-giro-hélice. El resto de los aminoácidos que componen el módulo CTPR permiten mutaciones sin alterar la estructura de estas proteínas de repetición. Además, se ha demostrado que las proteínas CTPR son más estables que sus homólogos naturales.

Los dominios de diseño TRAP son módulos evolucionados de dominios de reconocimiento TPR naturales. Diferentes módulos TRAP están diseñados para el reconocimiento de diferentes péptidos. Cada módulo TRAP promueve el reconocimiento de un péptido con una secuencia específica de 5 aminoácidos. Cabe destacar que los módulos TRAP no sólo permiten el reconocimiento de péptidos específicos con una reactividad cruzada mínima entre los diferentes módulos TRAP, sino que esta especificidad también se mantiene frente a otras proteínas celulares.

La estructura y estabilidad de los módulos CTPR, SCAB y TRAP están muy bien caracterizadas. Asimismo, las características de CTPR, SCAB y TRAP permiten utilizar estas proteínas en diversas aplicaciones, como andamiajes para la organización de enzimas en el caso de esta tesis.

En esta tesis se propone: 1) diseñar los módulos de ensamblaje TPR para generar ensamblajes con propiedades de ortogonalidad; 2) demostrar la versatilidad de los andamiajes proteicos sintéticos para la organización enzimática; 3) lograr un posicionamiento preciso a nanoescala de las enzimas; 4) desentrañar las cuestiones fundamentales relacionadas con el origen del mayor rendimiento de las enzimas ensambladas sobre andamiajes proteicos; y 5) generar pruebas de concepto que demuestren que los sistemas multi-enzimáticos organizados son eficientes y de interés científico-técnico e industrial. Para abordar estos retos, utilizamos un conjunto particular de proteínas de repetición diseñadas mediante ingeniería de proteínas.

Esta tesis consta de cinco capítulos, de los cuales los Capítulos 4 y 5 describen los resultados experimentales. En estos capítulos la combinación de métodos de ingeniería de proteínas y biocatálisis dio lugar a sistemas multi-enzimáticos organizados con un control nanométrico preciso de la organización espacial y las propiedades fisicoquímicas.

El **Capítulo 1** introduce el mundo de la biocatálisis haciendo hincapié en las ventajas de la inmovilización de enzimas en soportes sólidos frente al uso de enzimas libres. También se introducen las proteínas sintéticas y las propiedades que las hacen un excelente soporte para la organización de sistemas multicomponentes, como cascadas multi-enzimáticas y el estado actual de esta área de investigación. Por último, se introducen con detalle las diferentes proteínas de repetición que se aplicarán en este trabajo, especialmente las dos variantes sintéticas de los dominios TPR; TRAP y CTPR.

El **Capítulo 2**, en el contexto de la temática de la tesis, se introducen las enzimas que se van a ensamblar en los andamiajes basados en proteínas de repetición y se explica el objetivo general de la tesis.

El **Capítulo 3** describe en detalle la metodología y los protocolos experimentales utilizados en esta tesis.

El **Capítulo 4** describe el estudio sistemático de dos ensamblajes supramoleculares basados en proteínas ingenierizadas. El objetivo de generar módulos para ensamblaje controlado ha sido posible gracias a la selección de una proteína con características únicas, es decir, un andamiaje proteico basado en una secuencia consenso de ingeniería simplificada (CTPR). Los pocos residuos conservados en el módulo CTPR (8/34) permiten la introducción de mutaciones que dotan a la proteína de las funcionalidades a la carta requeridas en diferentes aplicaciones. Además, se ha comprobado que los módulos CTPR son más estables que sus homólogos naturales. Los módulos CTPR modificados para actuar como bloques de ensamble para la generación de andamiajes proteicos se denominan SCAffolding Bio-bricks (SCABs).

Para alcanzar el objetivo expuesto en el Capítulo 4, examinamos inicialmente dos procesos de ensamblaje distintos dentro del mismo marco, con el fin de formar dos ensamblajes supramoleculares basados en las interacciones cabeza-cola que los módulos CTPR presentan de forma intrínseca. Para asegurar la estabilización de estas interacciones cabeza-cola, ideamos e incorporamos residuos de coordinación, concretamente cisteínas e histidinas, para generar en base a el módulo CTPR los módulos SCAB. Los módulos SCAB que contenían cisteínas en las posiciones N- y C-terminal de los mismos dan lugar al establecimiento de interacciones cabeza-cola y su estabilización mediante la formación un enlace disulfuro covalente reversibles a través de las dos cisteínas ingenierizadas. Del mismo modo, los módulos SCAB que contenían histidinas lograron el mismo efecto mediante una interfaz formada a través de la coordinación entre las histidinas y los iones de cobre (II), consiguiendo una estabilización mediada por la coordinación del metal (Figura R2).

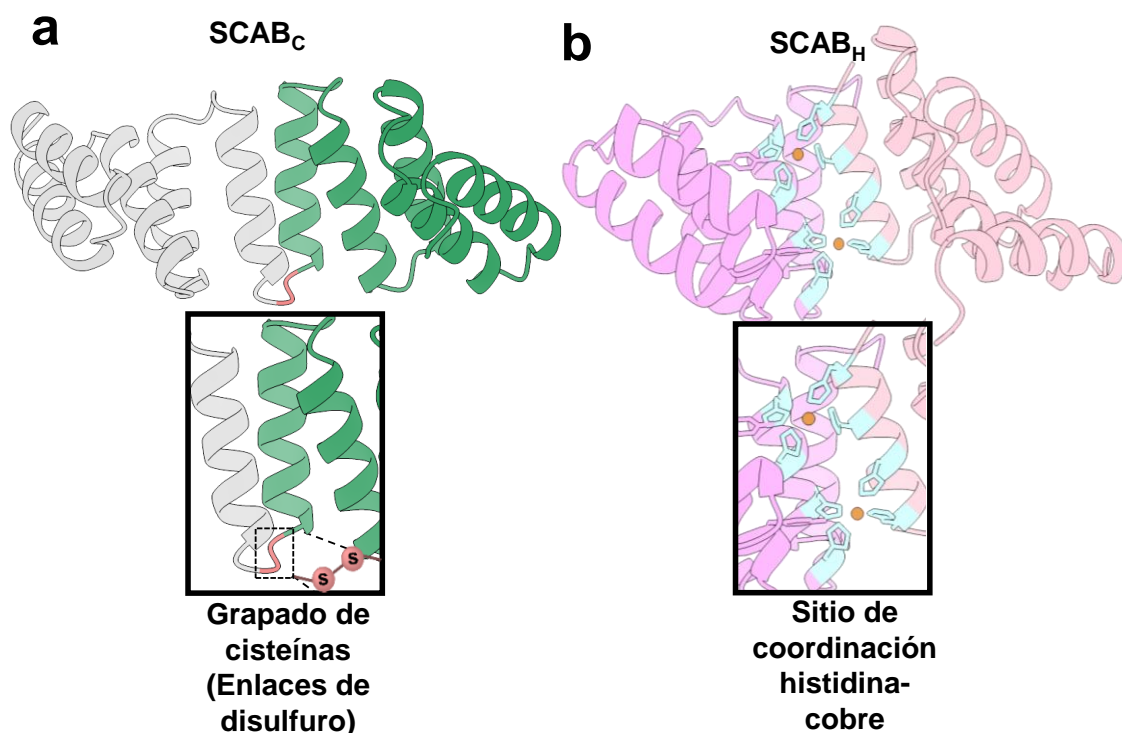


Figura R2. Representación de los dominios SCAB generados a través de la introducción de residuos de coordinación (cisteínas e histidinas) en los módulos proteicos CTPR. a) Esquema del ensamblaje de SCAB_{1C} y SCAB_{2C} estabilizado por la fijación de la interacción entre módulos a través un enlace disulfuro entre las dos cisteínas en el N- y C- terminal de los módulos dando lugar a SCAB_C (El enlace disulfuro entre las cisteínas ingenierizadas se muestra en rojo). b) Esquema del ensamblaje de SCAB_{1H} y SCAB_{2H} promovido por la coordinación específica histidina-cobre dando lugar a SCAB_H (la coordinación específica entre los residuos de diseño, histidinas, y cobre se muestra en azul claro, histidinas, y en marrón, cobre). Las estructuras 3D de los módulos SCAB se han modelado a partir de la estructura cristalina de la proteína CTPR8 (PDB ID: 2HYZ).

Una vez obtenidos los ensamblajes promovidos mediante un diseño previo de interacciones selectivas entre los módulos SCAB, se procedió a su caracterización, demostrando así la estructura y la estequiometría del ensamblaje. A continuación, se utilizaron estos ensamblajes prediseñados para promover la organización enzimática. Este objetivo se logró fusionando enzimas seleccionadas a módulos SCAB. Finalmente, se estudió la actividad enzimática y se determinaron los parámetros cinéticos de las fusiones de enzimas-SCAB libres y ensambladas.

Por último, se realizaron ensayos de acoplamiento enzimático para estudiar el efecto del andamiaje. El sistema multi-enzimático organizado consistente en dos enzimas ensambladas por anclaje a través de puentes disulfuro fue 3,5 veces más eficiente que las enzimas libres en relación con la producción del aminoácido natural L-Alanina en la biotransformación de relevancia industrial estudiada. Sin embargo, el sistema multi-enzimático organizado formado por dos enzimas y ensamblado por la coordinación específica entre histidinas y cobre no mostró diferencias en comparación con las enzimas libres en relación con la producción del aminoácido natural L-Alanina obtenida.

Como resultado, la investigación del Capítulo 4 sienta las bases para el desarrollo de dos nuevas metodologías de ensamblaje ortogonal basadas en proteínas CTPR para la creación de nuevos complejos multi-enzimáticos ordenados en la nanoescala. Ambas metodologías tienen su base en la ingeniería de proteínas y en el concepto de que la distancia entre enzimas desempeña un papel clave en la eficiencia catalítica.

El **Capítulo 5** pretende estudiar sistemáticamente cómo afectan los distintos parámetros a la eficacia catalítica de un sistema de ensamblaje de proteínas versátil y bien caracterizado basado en proteínas TRAP. Este enfoque ha sido posible gracias a la selección de un andamio proteico con características únicas, basado en un módulo de diseño simple y estable que al ser modificado es capaz de codificar biorreconocimiento ortogonal frente a secuencias peptídicas cortas. Adicionalmente la naturaleza modular del andamio permite el acoplamiento de dichos módulos en andamiajes proteicos nanométricos multivalentes, inalcanzables con ensamblajes proteicos convencionales.

Para cumplir el objetivo expuesto en el Capítulo 5, se fusionaron a las enzimas etiquetas peptídicas que son reconocidos de forma específica por los módulos TRAP. Una vez fusionados los péptidos, se lleva a cabo el ensamblaje enzimático a través reconocimiento biomolecular del andamiaje proteico TRAP multivalente generando sistemas multi-enzimáticos organizados compuestos por dos o tres enzimas (Figura R3).

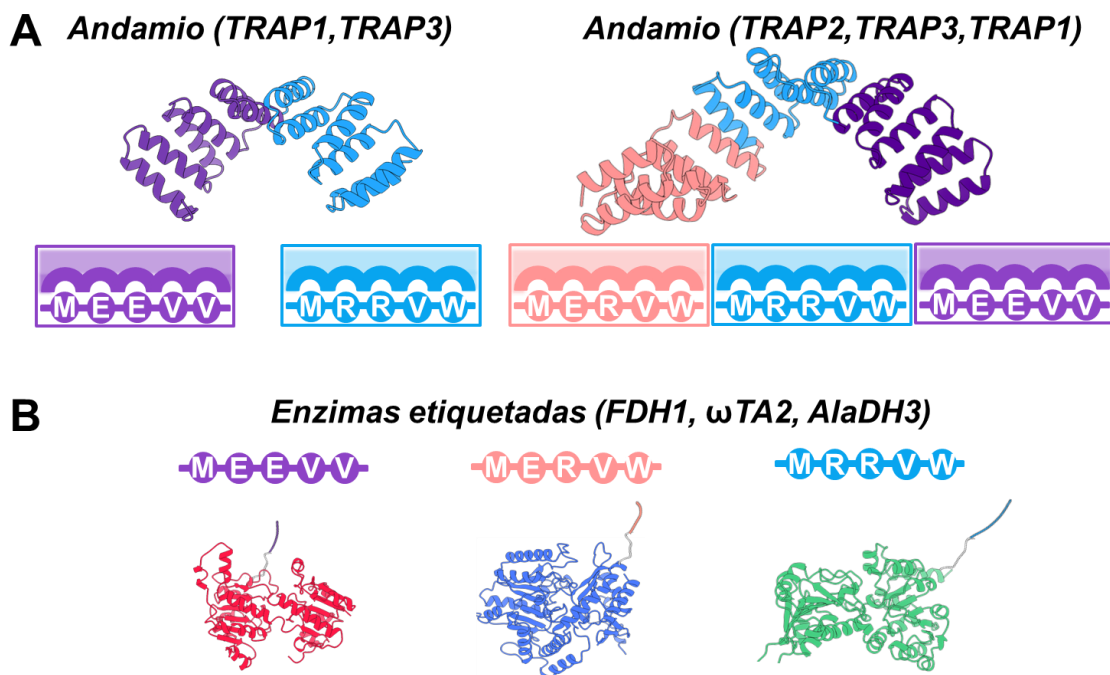


Figura R3. Esquema representativo del ensamblaje a través de reconocimiento biomolecular. a) Representación esquemática de los andamiajes basados en la proteína TRAP y sus péptidos afines y específicos compuestos por TRAP1-MEEVV (en morado), TRAP2-MERVW (en rosa) y TRAP3-MRRVW (en azul). b) Representación esquemática de los péptidos de reconocimiento afines a los módulos TRAP (MEEVV para TRAP1, MERVW para TRAP2 y MRRVW para TRAP3) fusionados a monómeros enzimáticos una FDH (en rojo, PDB ID: 5DNA), ω TA (en azul, PDB ID: 5LH9) y AlaDH (en verde, PDB ID: 1PJB). El sistema multi-enzimático organizado de dos enzimas consiste en el par FDH/AlaDH y el sistema de tres enzimas consiste en el trío FDH/ ω TA/AlaDH.

Una vez obtenidos los sistemas enzimáticos ensamblados sobre los andamiajes TRAP, estos se caracterizaron en detalle mediante técnicas espectroscópicas avanzadas que permitieron conocer la estructura, estequiometría y disposición de las enzimas sobre los andamiajes. A continuación, se estudió comparativamente la actividad enzimática y los parámetros cinéticos de las enzimas libres y ensambladas en los andamiajes TRAP.

Por último, se realizaron ensayos de acoplamiento enzimático para estudiar el efecto del andamio proteico TRAP en la actividad catalítica, así como para elucidar los mecanismos detrás de estos efectos. Los sistemas multi-enzimáticos organizados compuestos por dos y tres enzimas mostraron una actividad significativamente superior a la de las enzimas libres en dos biotransformaciones de relevancia industrial. En concreto se detectó un aumento de 5 veces para el sistema de dos enzimas y de 4 veces para el sistema de tres enzimas en relación con la producción del aminoácido natural L-Alanina y la molécula orgánica bencilamina, respectivamente. Adicionalmente, el sistema multi-enzimático de dos enzimas demostró ser más eficiente en la formación del aminoácido natural L-Alanina cuando el conjunto se inmovilizó en soportes sólidos para crear biocatalizadores heterogéneos. Nuestra investigación reveló que el incremento de actividad catalítica no sólo se debe al confinamiento espacial en el

andamio TRAP, sino también a la influencia de éste en las interacciones electrostáticas entre cofactores e intermediarios de la reacción.

Por tanto, el trabajo del Capítulo 5 arroja luz sobre el papel de la distancia entre las enzimas, pero también sobre las propiedades fisicoquímicas del ensamblaje basado en la proteína TRAP, y la mejora final del rendimiento catalítico de los sistemas ensamblados y la regeneración del cofactor NADH.

En conclusión, las diferentes tecnologías desarrolladas en esta tesis son generales y fácilmente extensibles a otros sistemas y pueden contribuir a la generación de sistemas multi-enzimáticos más robustos y organizados con precisión con aplicaciones industriales o medioambientales, así como a otras aplicaciones que requieran el control preciso en el espacio de diferentes elementos funcionales.

Summary

Biocatalysis is the scientific field related to the use of enzymes as catalysts in many important chemical conversions in industry. Enzymes present extraordinary advantages compared to their direct competitors, *i.e.*, chemical catalysts, such as higher catalytic performance, selectivity, and specificity. These properties make them ideal candidates for catalyzing reactions in the industry that cannot be carried out with conventional catalysts, such as complete conversion of substrates in a single step. However, the main drawback of enzymes, in addition to their high cost, is their susceptibility to denaturation when exposed to adverse conditions, such as high temperatures. This denaturation reduces their operational stability, affecting their applicability and causing a decrease in their catalytic activity.

In recent years, efforts have been made to address the challenges associated with enzyme immobilization by exploring the advantages of immobilizing enzymes on supports, specifically focusing on the spatial organization of enzymes. Although immobilization on solid supports is a robust method, it has significant drawbacks, such as the difficulty of co-immobilizing numerous enzymes, the slow chemical slowness required, or the lack of control in organization at submicromolar scales.

However, in the last decade, more modern enzyme support technologies have gained popularity, such as the use of biomacromolecules like DNA, peptides, nanoparticles, and proteins. The advantages of these supports are the ability to immobilize various enzymes, spatial control at different size scales, control of the interaction of substrates and cofactors with the supported systems, and better understanding and control of the reactions taking place in the enzyme cascade. By achieving a specific spatial organization of enzymes at the nanometer level, biomacromolecule-based supports not only provide spatial control but also improve catalytic efficiency and enable substrate and cofactor channeling effects.

In this thesis, in addition to addressing the previously described limitations of solid supports, we have gone a step further by using synthetic designer protein scaffolds as enzyme supports. These scaffolds based on TPR (Tetratricopeptide repeat) designer proteins have been engineered to allow the ordered nanoscale assembly of enzymes. The main features of these scaffolds are their modularity and their simplified and known structure that allow their modification without compromising their structure, as well as their intrinsic supramolecular self-assembly properties that enable them to assemble and provide precise spatial control at the nanoscale. These systems, therefore, outperform conventional protein scaffolds in advantages. Finally, these protein scaffolds, due to their ability to engineer their properties, can encode controlled interactions with essential cofactors, substrates, or reaction intermediates, which makes them very promising for various applications in the field of biocatalysis.

The two selected approaches to organizing enzymes in nanoscale space are supramolecular assembly through intrinsic self-assembly properties into linear structures of the selected scaffolds and biomolecular recognition-based assembly. Stabilization of the designed CTPR (Consensus Tetratricopeptide repeat) modules denoted as SCAB (SCAffolding Bio-Bricks) and assembled via inter-molecular head-tail interactions is achieved by implementing cysteine-mediated disulfide bonds and metal-directed assembly. Finally, for assembly by biomolecular recognition, we chose TRAP (Tetratricopeptide Repeat Affinity Protein) modules with a specific and orthogonal recognition capability against their target peptides (Figure S1).

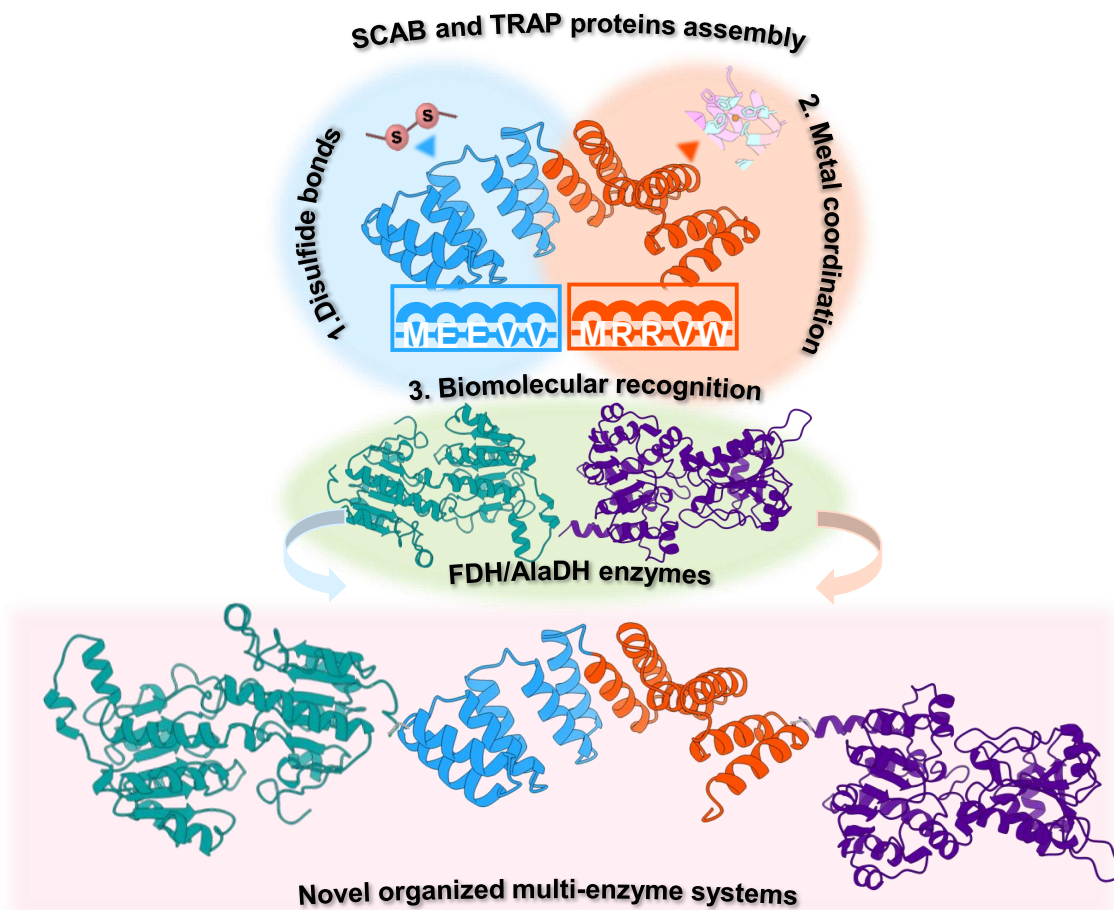


Figure S1. Schematic of novel organized multi-enzyme systems based on SCAB and TRAP proteins. Three enzymatic assembly strategies: SCAB self-assembly modules attached through cysteine-mediated disulfide bonds (1) or through metal coordination (2) and TRAP modules for assembly through biomolecular recognition (3).

Specifically, the protein scaffolds used in this thesis are two synthetic variants originating from the natural tetratricopeptide repeat (TPR) protein domain: self-assembly domains based on the consensus TPR sequence (CTPR) and designed to stabilize self-assemblies by different pathways (SCAB) and engineered domains with biorecognition properties (TRAP).

The consensus CTPR sequence is composed of 34 amino acids of which only 8 are highly conserved in sequence and define its characteristic folding giving rise to a helix-turn-helix structure. The remaining amino acids that make up the CTPR module allow mutations without altering the structure of these repeat proteins. In addition, CTPR proteins have been shown to be more stable than their natural counterparts.

TRAP-engineered domains are evolved modules of natural TPR recognition domains. Different TRAP modules are designed for the recognition of different peptides. Each TRAP module promotes the recognition of a peptide with a specific 5 amino acid sequence. Importantly, the TRAP modules not only allow the recognition of specific peptides with minimal cross-reactivity between the different TRAP modules but this specificity is also maintained against other cellular proteins.

The structure and stability of the CTPR, SCAB, and TRAP modules are well characterized. Also, the characteristics of CTPR, SCAB, and TRAP allow the use of these proteins in various applications, such as scaffolds for enzyme organization in the case of this thesis.

In this thesis, we propose to 1) design TPR assembly modules to generate assemblies with orthogonality properties; 2) demonstrate the versatility of synthetic protein scaffolds for enzyme organization; 3) achieve precise nanoscale positioning of enzymes; 4) unravel fundamental questions related to the origin of the enhanced performance of enzymes assembled on protein scaffolds; and 5) generate proof-of-concept demonstrating that organized multi-enzyme systems are efficient and of scientific-technical and industrial interest. To address these challenges, we use a particular set of protein-engineered repeat proteins.

This thesis consists of five chapters, of which Chapters 4 and 5 describe the experimental results. In these chapters, the combination of protein engineering and biocatalysis methods resulted in organized multi-enzyme systems with precise nanometric control of spatial organization and physicochemical properties.

Chapter 1 introduces the world of biocatalysis with an emphasis on the advantages of enzyme immobilization on solid supports versus the use of free enzymes. It also introduces synthetic proteins and the properties that make them excellent support for the organization of multicomponent systems, such as multi-enzyme cascades, and the current status of this area of research. Finally, the different repeat proteins that will be applied in this work are introduced in detail, especially the two synthetic variants of the TPR domains: TRAP and CTPR.

Chapter 2, in the context of the subject matter of the thesis, introduces the enzymes to be assembled into the repeat protein-based scaffolds and explains the overall objective of the thesis.

Chapter 3 describes in detail the methodology and experimental protocols used in this thesis.

Chapter 4 describes the systematic study of two supramolecular assemblies based on engineered proteins. The goal of generating modules for controlled assembly has been made possible by the selection of a protein with unique features, i.e., a protein scaffold based on a simplified engineered consensus sequence (CTPR). The few conserved residues in the CTPR module (8/34) allow the introduction of mutations that endow the protein with the à la carte functionalities required in different applications. In addition, CTPR modules have been found to be more stable than their natural counterparts. CTPR modules modified to act as assembly blocks for the generation of protein scaffolds are termed SCAffolding Bio-bricks (SCABs).

To achieve the goal outlined in Chapter 4, we initially examined two distinct assembly processes within the same framework to form two supramolecular assemblies based on the head-tail interactions that CTPR modules intrinsically exhibit. To ensure the stabilization of these head-tail interactions, we devised and incorporated coordination residues, namely cysteines, and histidines, to generate SCAB modules based on the CTPR module. The SCAB modules containing cysteines at the N- and C-terminal positions of the modules lead to the establishment of head-tail interactions and their stabilization by

Summary

forming a reversible covalent disulfide bond through the two engineered cysteines. Similarly, SCAB modules containing histidines achieved the same effect through an interface formed via coordination between histidines and copper (II) ions, achieving stabilization mediated by metal coordination (Figure S2).

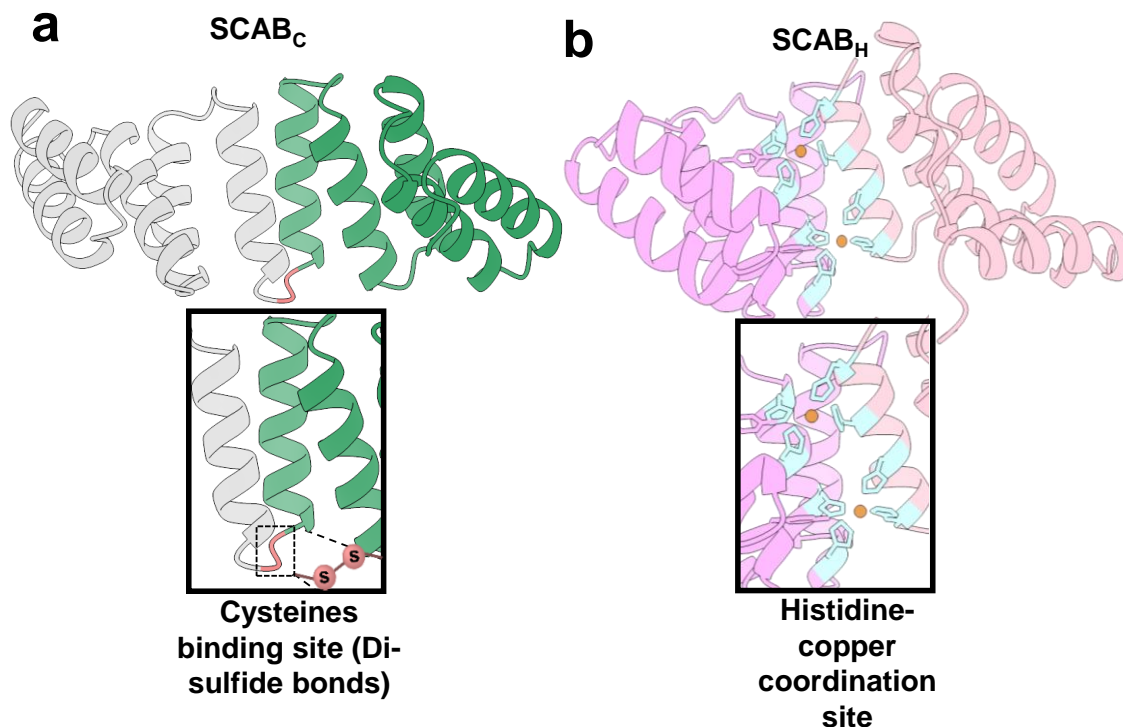


Figure S2. Representation of SCAB domains generated through the introduction of coordination residues (cysteines and histidines) into CTPR protein modules. a) Schematic of the assembly of SCAB1C and SCAB2C stabilized by the fixation of the inter-module interaction through a disulfide bond between the two cysteines at the N- and C-terminus of the modules giving rise to SCABC (The disulfide bond between the engineered cysteines is shown in red). b) Schematic of the assembly of SCAB1H and SCAB2H promoted by the specific histidine-copper coordination leading to SCABH (the specific coordination between the engineered residues, histidines, and copper is shown in light blue, histidines, and in brown, copper). The 3D structures of the SCAB modules have been modeled from the crystal structure of the CTPR8 protein (PDB ID: 2HYZ).

Once the assemblies promoted by a previous design of selective interactions between the SCAB modules were obtained, they were characterized, thus demonstrating the structure and stoichiometry of the assembly. These pre-designed assemblies were then used to promote enzyme organization. This goal was achieved by fusing selected enzymes to SCAB modules. Finally, enzymatic activity was studied, and kinetic parameters of free and assembled enzyme-SCAB fusions were determined.

Finally, enzyme coupling assays were performed to study the effect of scaffolding. The organized multi-enzyme system consisting of two enzymes assembled by anchoring through disulfide bridges was 3.5 times more efficient than free enzymes with respect to

the production of the natural amino acid L-Alanine in the studied industrially relevant biotransformation. However, the organized multi-enzyme system formed by two enzymes and assembled by the specific coordination between histidines and copper showed no difference compared to free enzymes in relation to the production of the natural amino acid L-Alanine obtained.

As a result, the research in Chapter 4 lays the foundation for the development of two new orthogonal assembly methodologies based on CTPR proteins for the creation of newly ordered multi-enzyme complexes at the nanoscale. Both methodologies have their basis in protein engineering and the concept that the distance between enzymes plays a key role in catalytic efficiency.

Chapter 5 aims to systematically study how different parameters affect the catalytic efficiency of a versatile and well-characterized protein assembly system based on TRAP proteins. This approach has been made possible by the selection of a protein scaffold with unique features, based on a simple and stable module design that when modified is able to encode orthogonal biorecognition against short peptide sequences. Additionally, the modular nature of the scaffold allows the coupling of such modules into multivalent nanometric protein scaffolds, unattainable with conventional protein assemblies.

To meet the objective outlined in Chapter 5, peptide tags that are specifically recognized by the TRAP modules were fused to the enzymes. Once the peptides are fused, enzyme assembly is carried out through biomolecular recognition of the multivalent TRAP protein scaffold generating organized multi-enzyme systems composed of two or three enzymes (Figure S3).

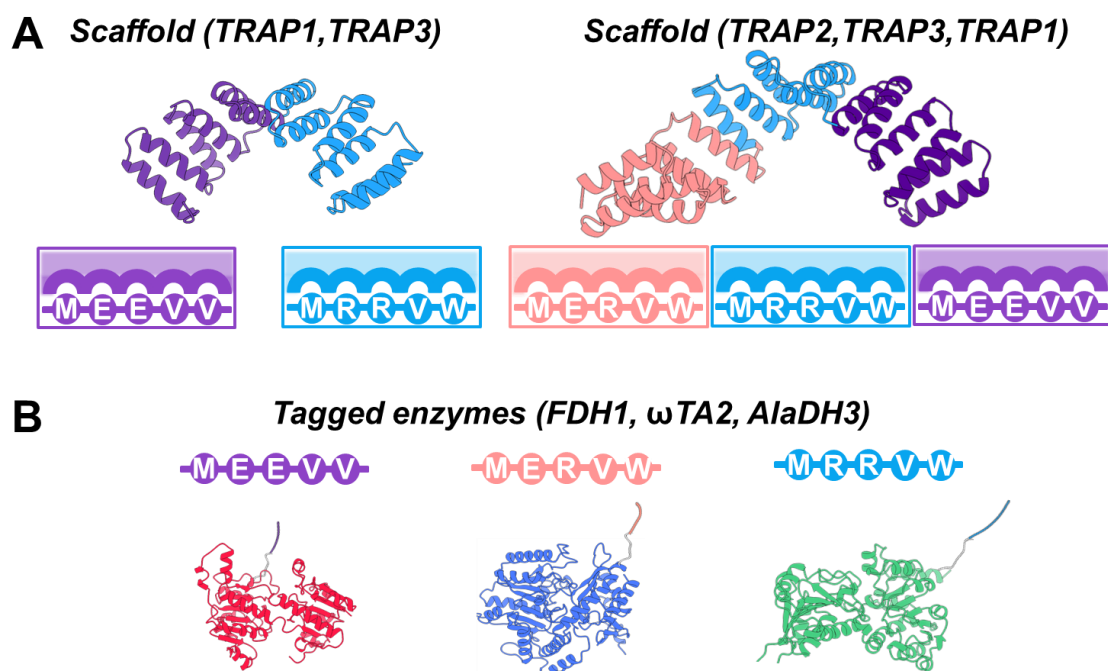


Figure S3. Representative schematic of assembly through biomolecular recognition. a) Schematic representation of TRAP protein-based scaffolds and their cognate and specific peptides composed of TRAP1-MEEVV (in purple), TRAP2-MERVW (in pink), and TRAP3-MRRVW (in blue). b) Schematic representation of recognition

Summary

peptides affine to TRAP modules (MEEVV for TRAP1, MERWV for TRAP2, and MRRVW for TRAP3) fused to enzymatic monomers FDH (in red, PDB ID: 5DNA), ω TA (in blue, PDB ID: 5LH9) and AlaDH (in green, PDB ID: 1PJB). The organized two-enzyme multi-enzyme system consists of the FDH/AlaDH pair, and the three-enzyme system consists of the FDH/ ω TA/AlaDH trio.

Once the enzyme systems assembled on TRAP scaffolds were obtained, they were characterized in detail by advanced spectroscopic techniques that provided insight into the structure, stoichiometry, and arrangement of the enzymes on the scaffolds. Then, the enzymatic activity and kinetic parameters of free and assembled enzymes on TRAP scaffolds were comparatively studied.

Finally, enzyme docking assays were performed to study the effect of the TRAP protein scaffold on catalytic activity, as well as to elucidate the mechanisms behind these effects. Organized multi-enzyme systems composed of two and three enzymes showed significantly higher activity than free enzymes in two industrially relevant biotransformations. Specifically, a 5-fold increase was detected for the two-enzyme system and a 4-fold increase for the three-enzyme system in relation to the production of the natural amino acid L-Alanine and the organic molecule benzylamine, respectively.

Additionally, the two-enzyme multi-enzyme system proved to be more efficient in the formation of the natural amino acid L-Alanine when the assembly was immobilized on solid supports to create heterogeneous biocatalysts. Our research revealed that the increased catalytic activity is not only due to spatial confinement in the TRAP scaffold but also to the influence of the TRAP scaffold on electrostatic interactions between cofactors and reaction intermediates.

Therefore, the work in Chapter 5 sheds light on the role of the distance between enzymes, but also on the physicochemical properties of the TRAP protein-based assembly, and the ultimate improvement of the catalytic performance of the assembled systems and the regeneration of the cofactor NADH.

In conclusion, the different technologies developed in this thesis are general and easily extensible to other systems and can contribute to the generation of more robust and precisely organized multi-enzyme systems with industrial or environmental applications, as well as to other applications requiring precise control in the space of different functional elements.

Chapter 1. General introduction

1.1. Biocatalysis

Catalysts are materials that reduce the energy barriers of a chemical transformation, thus accelerating the chemical process, without being consumed during the process ¹. Enzymes are vital biological catalysts in nature that execute both catabolic and anabolic processes, building cell metabolism that allows life to exist ². Compared to chemical catalysts, these protein-based catalysts, commonly named biocatalysts, are more environmentally friendly, as they are biodegradable and do not generate additional waste. In addition, they catalyze a wide range of chemical reactions, many of which are challenging to carry out. Moreover, enzyme-catalyzed reactions have more regio-, stereo-, and chemo-selectivity compared to chemical catalysts. Finally, many of them are biocompatible because they are found naturally in cells and catalyze reactions under soft conditions (pH and temperature) ² (Figure 1.1).

Biocatalysts are becoming a viable and competitive substitute for conventional chemical catalysts because they facilitate more environmentally friendly processes. This is because enzyme-catalyzed chemical synthesis helps develop chemical reactions based on cleaner, faster, and safer chemical processes. Given the irruption of enzymes in the biotechnology sector, it is essential to highlight the achievements and importance of enzyme technology in recent years. As a result, cell-free synthetic biology is developing as a compelling alternative to traditional whole-cell biology. This is because isolated enzymes have no genetic regulatory restrictions, and increasing chemical fluxes has no influence on the system's survival ³. However, it raises significant challenges in terms of both process- and cost-efficiency because these soluble systems frequently present low chemical yields, are noticeably unstable, and have a restricted re-usability ⁴ (Figure 1.1).

Enzymes have been used as biocatalysts in biotechnology for millennia to generate specific foods through whole-cell fermentations ⁵ and they were introduced to the broad public of our society as additives in commercial detergents ⁶. On the contrary, the success of industrial enzymes began towards the end of the previous century with the introduction of enzyme-guided evolution. These advancements have fuelled the use of isolated enzymes in the synthesis of non-natural chemical compounds, including the biomanufacturing of pharmacological intermediates, and their use in medical therapy ⁷ (Figure 1.1).

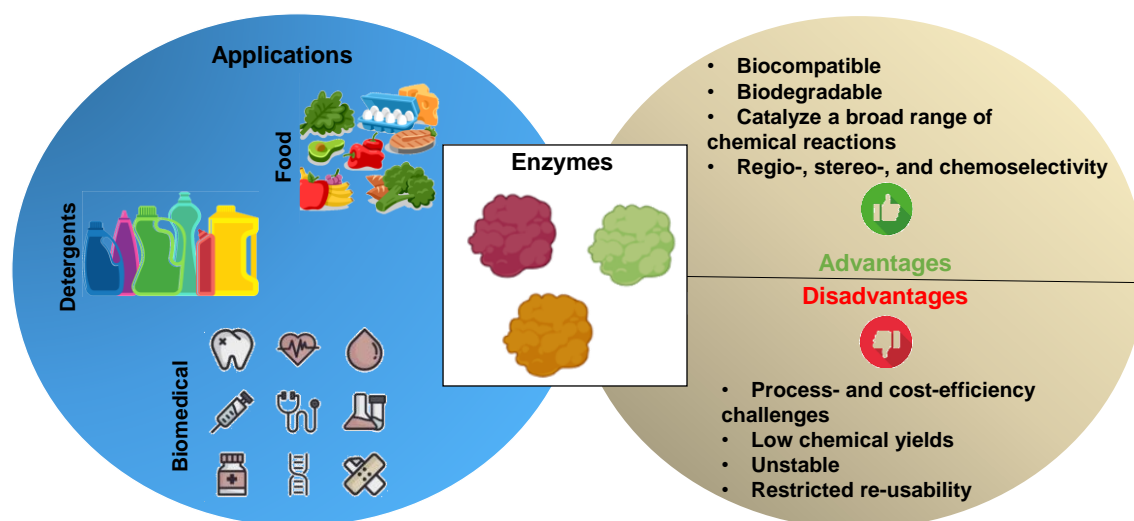


Figure 1.1. The star of catalysts: Enzymes. These protein-based catalysts (biocatalysts) have multiple applications, from industrial applications such as the food industry or detergents to biomedical applications, such as pharmacology. The most typical advantages and disadvantages of enzymes are compiled.

Within the field of biocatalysis, multi-enzyme systems are rising interest among scientists because they selectively catalyze chemical reactions under moderate circumstances, boosting the sustainability of chemical processes⁸. Stepwise biocatalysis in a single pot has substantial benefits over traditional step-by-step techniques reducing purification processes and improving overall productivity and efficiency. In this context, synthetic biology and metabolic engineering have demonstrated tremendous promise in meeting the demand for more efficient and sustainable chemical production methods⁹. However, owing to the inherent limits of cells, *in vivo*, techniques confront significant obstacles^{8,10}. In addition to *in vivo* systems, isolated enzymes may be used *ex-vivo* to catalyze both natural and non-natural tandem processes, resulting in a novel concept: systems biocatalysis. This new field explores the simplest biology to create the most complicated chemistry¹¹. Multi-enzyme systems can catalyze a high number of progressive reactions in one pot, avoiding the build-up of intermediates (toxic or inhibitory), reusing the system cofactors more effectively, and thereby decreasing waste formation while enhancing product yield. Additionally, isolating enzymes from their living environment results in the loss of spatial compartmentalization, which improves the catalytic efficiency of biosynthetic pathways. To improve the efficiency of *in vitro* multi-enzyme systems, novel approaches to obtain proper control over the local concentration of reactants and intermediates are required^{12,13}. The use of heterogeneous biocatalysts (catalysis where the phase of catalysts differs from that of the reactants or products) is required to enable enzyme recycling and achieve profitable processes economically viable^{14,15}. Since enzymes are often expensive catalysts, it is often necessary to recover and reuse them to reduce the costs associated with their use.

In recent years, significant efforts have been made to overcome the limitations associated with free enzymes in multi-enzyme systems. These efforts compile a range of strategies, including enzyme re-engineering through direct evolution, semi-rational design, and rational design. Additionally, techniques like enzyme fusion using tiny linkers and enzyme immobilization using diverse materials and procedures have been explored as potential solutions.

Enzyme re-engineering encompasses different approaches to enhance enzyme properties. Direct evolution involves introducing genetic diversity through random mutagenesis or DNA shuffling and selecting improved enzyme variants. Semi-rational design modifies enzymes based on known structure-function relationships to optimize catalytic activity and compatibility. The rational design employs computational tools to precisely modify enzymes, enhancing stability and efficiency (Figure 1.2a). The semi-rational design has been recognized as a powerful tool for protein engineering, although further understanding of structure-function correlations is needed (Markel et al.,¹⁶). Directed evolution, on the other hand, utilizes iterative rounds of diversity creation and screening to improve enzymes, even with limited structural understanding. With the availability of gene synthesis, sequencing, and bioinformatic tools, directed evolution is widely applicable in laboratories worldwide. Computationally assisted semi-rational design

plays an increasingly important role in evolving natural enzymes and can even enable the creation of catalytically active proteins from scratch ^{16,17}.

Enzyme fusion involves physically linking enzymes together, which facilitates efficient communication and enhances catalytic efficiency (Figure 1.2b). Chen et al. ¹⁸ have highlighted the formation of multi-enzyme systems through recombinant fusion proteins, which often require component proteins and linkers. These linkers, like natural multi-domain proteins, connect functional domains and offer additional benefits. They not only bind the protein moieties but also contribute valuable features like preserving cooperative inter-domain connections ¹⁹ and biological activity. Investigations into linkers in natural multi-domain proteins have generated hypotheses about protein fusion ¹⁸. Empirical linkers with varied sequences and conformations have been developed for the construction of recombinant fusion proteins. Flexible linkers, rigid linkers, and cleavable linkers are three common types that have demonstrated diverse applications in fusion protein assembly ¹⁸.

Enzyme immobilization onto solid supports provides stability, protection, and reusability, improving the overall performance of multi-enzyme systems (Figure 1.2c). This is particularly relevant in industrial biocatalysis, as enzyme recovery and reuse are essential. Enzyme immobilization ensures the recyclability of enzymes, allowing for their repeated use. However, successful immobilization requires enzymes with high stability or the incorporation of stabilizers to preserve the enzyme structure ¹⁵. While numerous immobilization procedures exist, there is ongoing research to develop novel protocols that can further enhance enzyme characteristics during immobilization, presenting an attractive goal ²⁰.

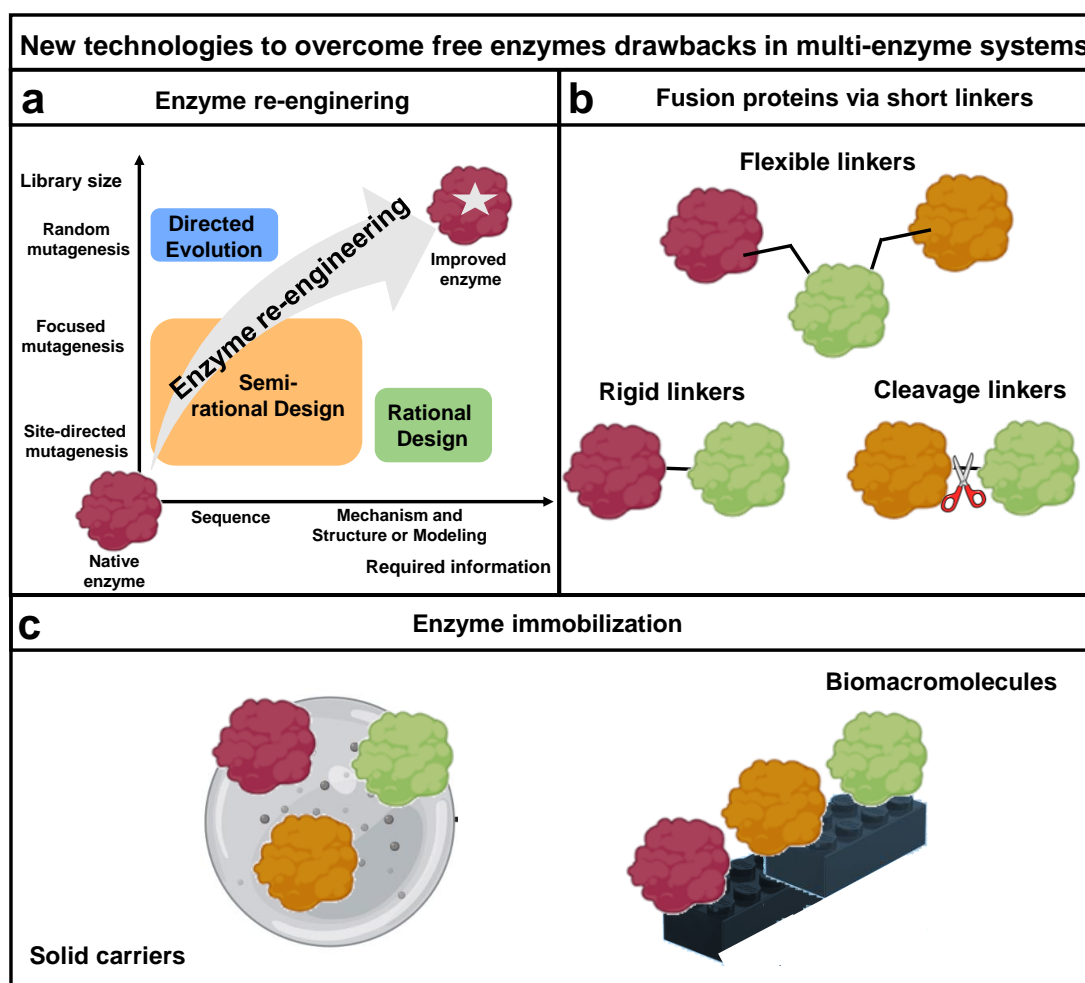


Figure 1.2. Schematic representation of new technologies to overcome the drawbacks of free enzymes in multi-enzyme systems. Among the most used new technologies are these three subgroups: Enzyme re-engineering, enzyme fusion via peptide linkers, and enzyme immobilization. a) Enzyme re-engineering evolved from direct evolution on semi-rational design and then on rational design. This process gives improved enzymes based on mutations and computational approaches²¹. b) Enzymes are easily fused in the presence of small linkers and are divided into three groups: flexible, rigid, and cleavage linkers¹⁸. c) Enzyme immobilization has evolved from traditional methods such as solid carriers to more advanced methods like the inclusion of biomacromolecules such as DNA, peptides, nanoparticles or even proteins.

1.2. Enzyme immobilization on supports

During the last half-century, enzyme immobilization defined as the confinement of one or even more enzymes in a specific area, was created to tackle the challenges of enzyme recovery, solubility, and reuse¹⁴. Immobilization is a very powerful method to overcome virtually all drawbacks of free enzymes if the right strategy is designed. Equally important is the fact that by immobilizing the enzyme, its inactivation induced by inactivating agents (temperature, pH, and organic solvents) is reduced, which allows them to be used for a longer time due to their greater stability; in addition to allowing their long-term storage. Moreover, by immobilizing enzymes, improved control over enzyme processes has been observed, which can be crucial in some instances. For these reasons, enzyme

immobilization has been the most extensively utilized strategy to overcome the limitations of free enzyme solubility and robustness ²².

Co-immobilization of enzymes is one of the most used procedures in enzyme immobilization. It has been previously reported by García-Galan et al. ⁴ that co-immobilization accelerates the initial response rate by reducing or even eliminating the delay time that occurs when numerous enzymes are immobilized on distinct particles or even free. This lag arises because the concentration of the intermediates is initially relatively low, preventing the other enzymes in the reaction chain from expressing their activity from the start of the process. Furthermore, when co-immobilized enzymes are used, the initial concentration of the intermediates generated in a limited location can be quite high, allowing the other enzymes to express all of the activity from the beginning of the process ^{12,23,24}.

1.2.1. Solid supports

The use of pre-existing materials (solid carriers) as an immobilization matrix is one of the most extensively utilized approaches for co-immobilizing enzymes (Figure 1.3). For enzyme co-immobilization, a wide range of commercial carriers with varying mechanical properties (flexibility, stiffness, pore size, etc.) are available and might easily provide quantities of hundreds of kg. These enzyme systems are referred to as the above-mentioned heterogeneous biocatalysts. As a result, the enzyme becomes insoluble in water in solid form during the immobilization process and shows different characteristics than homogeneous biocatalysts (soluble enzyme) which are dispersed in an aqueous-liquid phase ²⁵.

Co-immobilization of enzymes inside the porous structure of a solid carrier allows the enzyme molecules to be entirely solid and dispersed without the potential of interacting with any external agent. Thus, the enzymes are protected from contact with the crude enzyme extract, preventing aggregation, autolysis, or proteolysis by the extract proteases (which will also be dispersed and immobilized). Furthermore, the co-immobilized enzymes will not encounter any external hydrophobic surface, such as air bubbles generated by the delivery of some needed gases or by the vigorous agitation required to maintain pH. These gas bubbles can inactivate soluble proteins enzymatically ^{15,26}, but not enzymes co-immobilized on a porous substrate ²⁷.

Several factors are key when choosing solid carriers, for example, the carrier must have a large internal surface area to have excellent geometric congruence with the surface of the enzyme. Also, if the carrier consists of very thin fibers, such as those smaller than the protein, it will be difficult to create an intense connection between the enzymes and the carrier. Finally, the carrier must have a high density of reactive groups on its surface. This is key since intensive multipoint covalent binding can only occur when there are several reactive groups on the carrier below the surface of the protein. Moreover, the reactive groups on the carrier involved in enzyme co-immobilization must react with groups often found on the enzyme surface and they must be stable enough to enable for extended durations of enzyme-carrier reaction. Equally important the reactive groups of the protein and the carrier must provide a little steric barrier to the reaction because multipoint covalent bonding needs interaction between groups connected to stiff structures after initial enzyme co-immobilization. Additionally, after enzyme co-immobilization, it should be

simple to obtain an inert final surface on the carrier by removing or blocking the remaining reactive groups on the support without impacting the enzyme.¹⁵

Epoxy or glyoxyl groups are particularly appropriate among reactive groups²⁸. Glyoxyl agarose has the further benefit of directing immobilization via the area(s) of the protein-rich reactive residues, allowing for a strong multipoint enzyme-support reaction²⁹. Another typical method for co-immobilizing enzymes is via cross-linking with glutaraldehyde. In several situations, it has offered strong stability factors³⁰, as this method is quite adaptable and can be applied in a variety of ways³¹.

Zeballos et al.³² and Santiago-Arcos et al.³³ employed different immobilization chemistries to co-immobilize enzymes on solid substrates. The first combines protein engineering and solid carriers by immobilizing two enzymes on a protein-based scaffold, which is then immobilized on a solid carrier, an agarose porous microbead. Because of the inclusion of cobalt chelating groups with the proteins' his-tagged, the multi-enzyme system is immobilized on this solid carrier. However, to promote enzyme co-immobilization, the latter has gone a step further and produced a tri-heterofunctional solid carrier with three functionalization chemistries. This carrier is made up of agarose microbeads that have been activated with aldehyde, amino, and cobalt-chelate groups. In this work, they show that the tri-heterofunctional carrier can co-immobilize up to three enzymes.

Furthermore, when co-immobilizing enzymes, the following criteria must be taken into account from the point of view of biocatalyst design²³:

- The enzymes should be similar in size.
- The immobilization rate and loading capacity of the enzymes should be assessed individually.
- The immobilization approach for all enzymes must be the same.
- The different stabilities of co-immobilized enzymes should be studied.

Besides, multipoint covalent binding of enzymes on pre-existing highly activated supports through short spacer arms. This implies that numerous residues on the enzyme surface increase the structural rigidification of the co-immobilized enzymes and in rare situations, the stability of the enzyme may decrease after co-immobilization¹⁵. Even so, the most prominent limitations of enzyme co-immobilization are loss of stability, activity, specificity, and selectivity, and reduction of reaction inhibition. When enzymes are co-immobilized it is possible to achieve longer reaction times, a higher variety of reaction conditions, and higher yields^{4,14,15}. Also highlight that enzyme co-immobilization on solid carriers, can pose mass transfer issues, for example, if the carrier can establish unwanted contacts with the enzymes^{12,34}.

Among the solid supports considered highly promising, nanoparticle enzyme clusters emerge as a particularly exciting avenue. Nanoparticle-enzyme clusters involve immobilizing enzymes on nanoparticles, leading to enhanced stability, controlled arrangement, and improved catalytic performance. This approach offers benefits like increased enzyme activity, improved substrate accessibility, and the potential for multifunctional enzyme systems³⁵. Breger et al.³⁵ demonstrate that enzymes constituting a multistep cascade can self-assemble with nanoparticle scaffolds into nanoclusters that access substrate channeling and improve catalytic flux by orders of magnitude.

1.2.2. Biomacromolecule-based supports

In the last years, more complex techniques have evolved to address the limitations of enzyme co-immobilization on solid supports. For example, chemical induction-based and biomacromolecule-based enzyme co-immobilization.

A framework of chemical induction-based assemblies has been widely used, which may be split into three categories: inhibitors, cofactors, and metal ions³⁶. Inhibitors reduce enzyme activity by binding to them in a specific way, which is determined by their chemical structure and interactions with active sites on enzymes³⁶. For cofactor-dependent enzymes such as oxidoreductases and transferases, intramolecular immobilization of enzyme subunits by cofactor binding has been widely reported³⁶. Metal ions can drive proteins to form large assemblies by modulating non-covalent interactions mediated by metal ions, such as metalloenzymes. Metal ions are frequently coordinated by oxygen, nitrogen, or sulfur centers of amino acid residues at active sites, influencing enzyme activity and stability³⁷.

The biomacromolecule framework has been frequently used for enzyme co-immobilization and can be divided into three categories: DNA-, peptide-, and protein-induced assembly (Figure 1.3).

DNA origami, a form of DNA bioconjugation, enables enzyme self-organization by folding DNA strands into nanostructures³⁸. Enzymes can be placed and regulated by combining DNA strips with enzymes that bind covalently³⁹ and their attachment can be achieved through chemical conjugation like click chemistry⁴⁰. DNA origami allows for precise control over enzyme arrangement, stability, and protection, resulting in enhanced enzyme complexes with improved functionality³⁸. Klein et al.⁴¹ achieved a three-enzyme sequential cascade using DNA origami. They immobilized DNA-conjugated amylase, maltase, and glucokinase on a self-assembled DNA origami triangle. The catalytic enhancement observed was attributed to increased enzyme stability and a localized DNA surface affinity or hydration layer effect, rather than directed enzyme-to-enzyme channeling.

Peptide conjugation can be accomplished by gene fusion to insert a peptide linker or through enzyme-catalyzed peptide cross-linking. Within the multi-enzyme architecture, peptide linkers offer the benefit of linking and separating catalytic domains⁴². Unfortunately, as Wang et al.³⁶ reported the overexpression of enzymes linked to peptide linkers, on the other hand, frequently results in the creation of an inclusion body, which not only impairs enzymatic activity but also results in poor enzyme recovery.

The use of protein-based scaffolds having programmable features, in which both functional and scaffold domains may be genetically encoded and translated as a single chain is a promising strategy. In this sense, the utilization of tiny modular scaffold units whose assembly is dependent on certain well-defined and controllable interactions opens the door to a novel and easy technique for enzymatic component spatial layout⁴³. Zeballos et al.³² and Zhang et al.⁴⁴ report two types of protein-based scaffolds in which enzyme co-immobilization was successful. The first immobilizes two enzymes on a protein-based scaffold which, in turn, is immobilized on a solid carrier, a porous agarose microbead. Regarding the protein-based scaffold, the enzymes are co-immobilized by a cellulosome-

based system, the dockerin-cohesin pair (the basis of this protein-based scaffold is developed in section 1.4). The second immobilizes an enzyme by another widely used protein domain-based protein scaffold, the SpyTag-SpyCatcher system (the basis of this protein scaffold is developed in section 1.4).

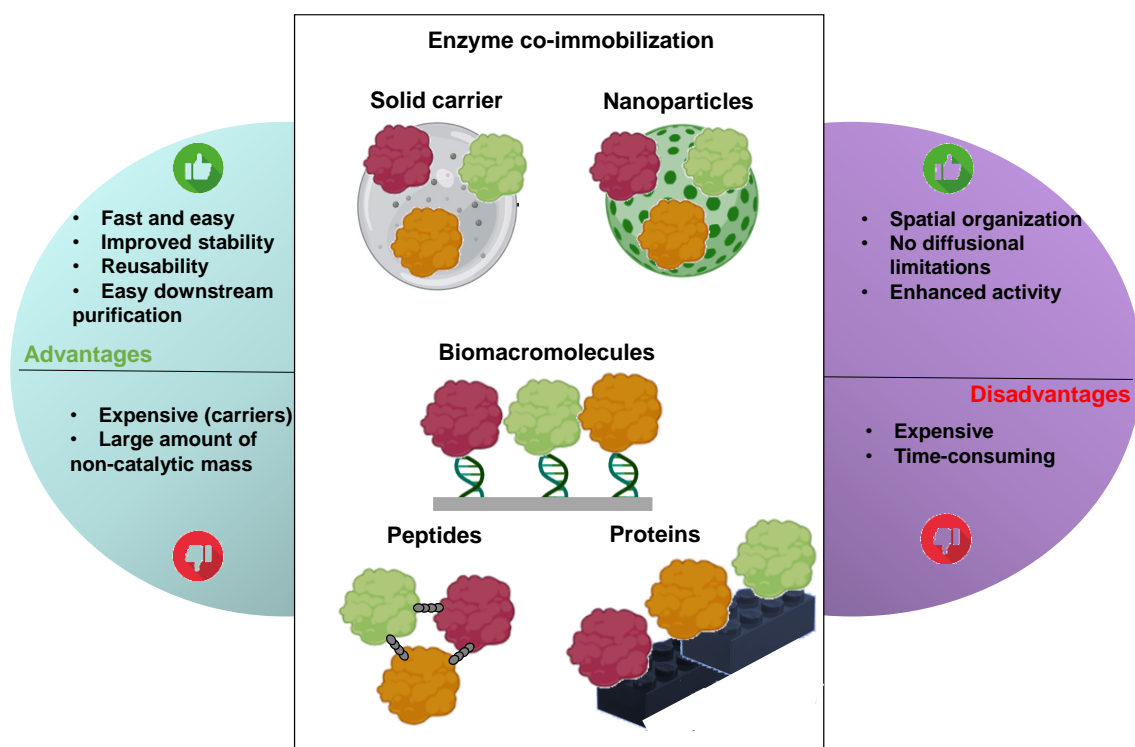


Figure 1.3. Schematic representation of enzyme co-immobilization by two alternative routes. The first consists of using a solid carrier or nanoparticles- on which the enzymes are co-immobilized by the presence of reactive groups and different chemistries. The second consists of using biomacromolecule-based scaffolds (DNA-, peptides-, and proteins) in which the enzymes will adopt a precise spatial organization. The most significant advantages and disadvantages of each of the co-immobilization routes are shown.

1.3. Proteins as building blocks for enzyme immobilization: general strategies, spatial organization, and substrate channeling effect

In the last decade, protein engineering has greatly enhanced the catalytic efficiency of several enzymes. In contrast, improving enzyme stability remains a challenge due to the lack of knowledge of the structure-function relationship, which hinders the discovery of suitable mutation sites. It is anticipated that protein stabilization can also be achieved through protein engineering, with a substantial impact on the development of new robust enzymes ²¹.

Having immobilized enzymes in protein scaffolds ^{42,43} has several benefits, including the ability to employ them both *in vivo* and *in vitro*, control over the spatial organization, and the elimination of diffusional restrictions ⁴⁵. Theoretical frameworks and analytical techniques to track the progress of the multi-enzyme system in terms of activity and stability, on the other hand, would still be needed. Other factors, such as the

microenvironment and the distance between enzymes, may also help boost catalytic efficiency^{3,46}. Protein-based scaffolds are found in nature and act as docking sites for various protein elements of the signaling cascade, facilitating the corresponding connections and activities⁴⁷.

Protein-based scaffold systems can be designed using a variety of techniques. The three fundamental structural components of protein-based scaffolds are the adapter domain, the peptide motifs/ligands, and the linker. The overall shape of the scaffold is influenced by these three building pieces⁴⁸. Small protein-binding modules known as adapter domains are found in adaptor proteins, which enable protein-protein interactions in a tightly controlled manner. Linearly interacting peptides with short amino acid sequences that complement the adapter domains are known as ligands or peptide motifs. The ligands serve as a link between the bound peptide motifs and the designed enzymes (Figure 1.4a). Alternatively, linkers can be omitted in the construction of scaffold systems by directly fusing peptide ligands to enzymes (Figure 1.4b). While linkers can be used to directly fuse enzymes to protein-based scaffolds for the creation of multi-enzyme complexes (Figure 1.4c) or to combine several enzymes into multi-enzyme complexes (Figure 1.4d)⁴⁹.

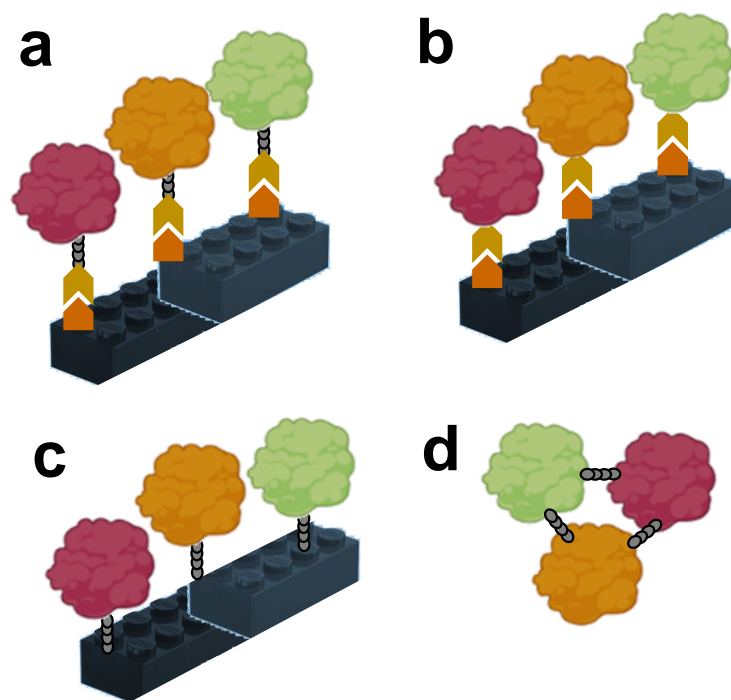


Figure 1.4. Schematic representation of different strategies for enzyme immobilization on protein-based scaffolds. a) Adapter domain-ligand-mediated assembly of the enzymes on the protein scaffold with a peptide linker serving as a connecting linker. In the adapter-ligand domain, the adapter is represented in orange and the ligand is represented in yellow. The peptide linkers are represented in grey. b) Direct assembly of enzymes on scaffold proteins through adapter domain-ligand interactions. In the adapter-ligand domain, the adapter is represented in orange and the ligand is represented in yellow. c) Direct fusing of enzymes on scaffold proteins via peptide linkers. d) Peptide linker-based enzyme cross-linking. The peptide linkers are represented in grey.

The protein scaffold is represented by two Lego blocks (in blue tones). The peptide linkers are represented in grey.

The elements of an enzyme cascade are precisely positioned at nanometer distances by co-immobilization of enzymes on protein-based scaffolds. As it was mentioned above, using a protein scaffold, in contrast to alternative supports, has recently been found to significantly enhanced performance inside these enzyme cascades, although the reason for this improvement in catalytic activity is still unclear. Experts in the field agree that the spatial organization of enzymes is one of the most compelling possibilities. In the protein-based scaffolds, enzymes acquire a precise spatial arrangement through scaffolding. Understanding how the local density and spatial arrangement of connected enzymes affect the diffusion of intermediates between enzymes and, consequently, the ultimate rate of product synthesis is fundamentally lacking⁵⁰⁻⁵². Furthermore, the multi-enzyme complex formed on a protein-based scaffold reduces product feedback inhibition, overall transit length, and intermediate loss caused by catalytic site closeness^{42,53}. Easier substrate transport is due to the presence of intermediates surrounding the active sites of scaffolded enzymes. Therefore, if the scaffold has additional units of an enzyme wrapped around the first enzyme, the effective concentration of intermediates will increase. Since a diffusion-limited trapping effect is seen, the first enzyme's intermediate will encounter a second enzyme before diffusing into the medium and will transform. As a result, the protein-based scaffold gets multiple units of the first enzyme to surround it, or a diffusion-limited trapping effect may take place. If either of these phenomena takes place, the intermediate of the first enzyme would reach a second enzyme. Finally, the intermediate would be transformed into the final product before diffusing into the medium, thus increasing the effective concentration of the intermediate (Figure 1.5).

Substrate channeling offers several benefits for metabolic pathways. Firstly, it enhances efficiency by reducing intermediate loss, minimizing side reactions, and increasing substrate conversion. Secondly, it protects intermediates from degradation or unwanted reactions by facilitating their direct transfer between enzymes. Thirdly, it prevents competitive reactions by ensuring that intermediates are not intercepted by other enzymes or molecules. Lastly, substrate channeling allows for enhanced regulation through the spatial organization, enabling fine-tuning of metabolic activity by controlling enzyme proximity and availability^{3,54}. Thus, the substrate channeling effect, achieved through the spatial organization of enzymes involved in sequential reactions, enhances metabolic pathway efficiency and specificity by directly transferring intermediates between active sites, minimizing diffusion and loss, and promoting their efficient utilization (Figure 1.5)^{3,24,55-57}.

Ellis et al.³ concluded that using protein-based scaffolds for multiple enzyme immobilization provides access to the previously mentioned potential functional benefits, which fall into three loosely grouped and related mechanisms, namely substrate channeling, enzyme sequentially, proximity, confinement, or some combination thereof, and localized scaffolding effects³. Another advantage of protein scaffolding is the claimed improvement in catalytic efficiency based on product formation. In the studied systems using this sort of protein-based scaffolds, the observed gains reach an increase of up to 3.6 in comparison with the free enzymes³.

Nanoscale protein-based scaffolds enable easy co-immobilization of complex enzymatic cascades, enhancing performance through substrate channeling and route flow. The genetically tuneable spatial organization of enzymes in protein-based supports⁵⁸ plays a crucial role in improving catalytic activity and biochemical reaction efficiency, with wide implications in biocatalysis, metabolic engineering, and synthetic biology.

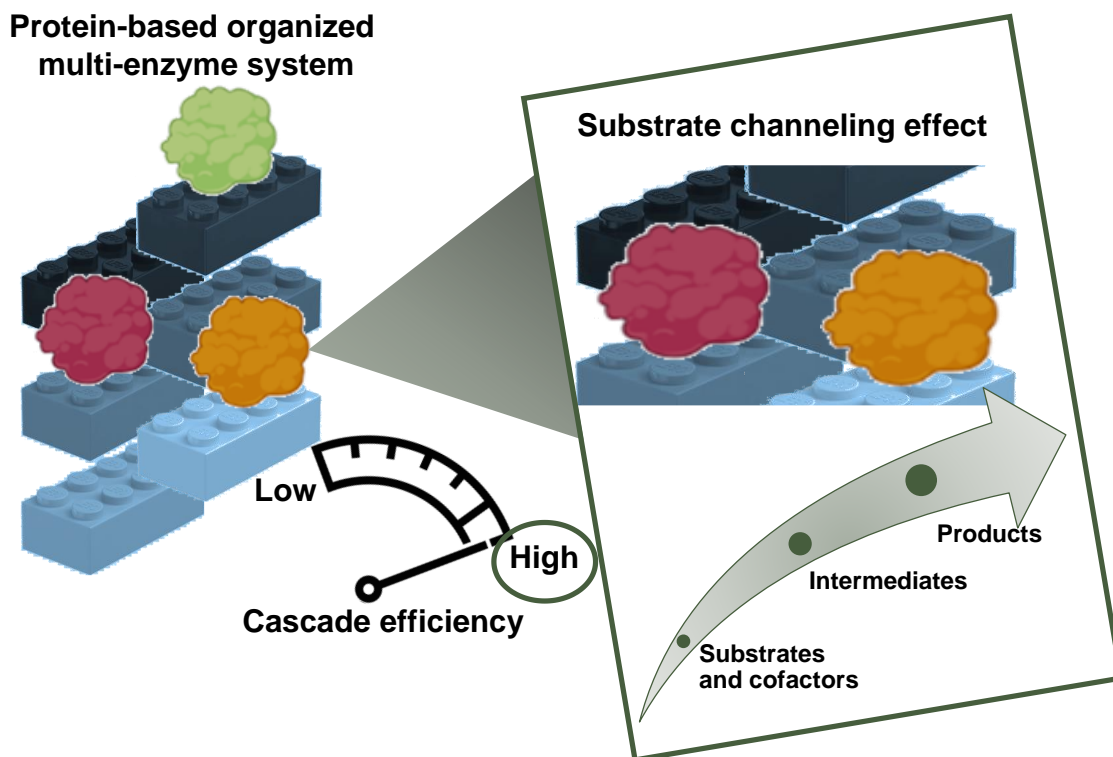


Figure 1.5. Substrate channeling effect on enzymes scaffolded on protein-based scaffolds. Co-immobilization of enzymes on protein-based scaffolds enables precise spatial structure, which improves cascade efficiency in terms of product production. The substrate channeling effect was performed on this multi-enzyme system, in which substrates and cofactors are directly transported from one enzyme to another without the loss of intermediates.

1.4. Enzyme immobilization onto protein-based scaffolds

Recent studies stated that the combination of protein engineering with enzyme immobilization is a promising tool to obtain novel systems for many applications, such as flow biocatalysis⁵⁹. Furthermore, the use of protein scaffolds is a promising approach alternative to try to mimic the frequent enzyme scaffolding observed in natural systems.

This combination aims not only to co-immobilize enzymes at the nanoscale to improve the effectiveness of catalytic cascades but also to act as a unique tool for the creation of sustainable methods for the manufacture of high-added value chemicals from renewable feedstocks. Moreover, as a result of this combination, it is possible to create novel and potent bionanomaterials for use in commercial catalytic bioprocesses⁶⁰.

Cohesin-dockerin interactions obtained from natural cellulosomes are the most widely used protein-based scaffolds for structuring multi-enzyme systems. The connection

between cohesin and dockerin is an example of a high-affinity protein-protein interaction (Figure 1.6a). This relationship is universal within a specie, but it is unique between species. As a result, scaffold-transported cohesin from a certain species should recognize and bind the majority or all enzyme-transported dockerins from that specie, but not dockerins from other species. Since the calcium-binding motif of the dockerin domain is similar to the helix-loop-helix structural domain of eukaryotic calcium-binding proteins, calcium ions are also significant in link ⁴⁹. For instance, dockerin domains located at the C-terminus of cellulosome enzyme components and cohesin repeat modules positioned on a non-catalytic molecular scaffold tightly connect to form cellulosomes ³². While it is true that nature offers a limited diversity of cohesin-dockerin pairs, these can be assembled as synthetic cellulosomes formed by fusion proteins containing different cohesin domains (scaffoldin) that reversibly bind to various enzymes fused to their cognate dockerin domain through calcium-driven protein-protein interactions with extremely high affinity (K_D in the nM-pM range) ⁴⁹. Hence, cohesin modules associate with scaffoldin and oversee organizing the cellulolytic enzymes, dockerin-cohesin interactions are a crucial component of the cellulosome ^{42,61}.

The protein scaffold created by Karpol et al. is one of the most illustrative examples. An affinity-based protein purification approach using dockerin-cohesin interactions, in which the target protein was tagged with shortened dockerin (affinity tag), and the cohesin module was immobilized on the bead cellulose (affinity resin matrix). Through cohesin-dockerin interactions, the target protein attaches to the column matrix and is successfully removed using EDTA gradients. The reusability investigation provides additional evidence of the recyclable nature of the affinity matrix for protein purification ⁶². To perform a variety of enzyme cascades more effectively, such as the hydrolysis of bio-based resources (such as pretreated cellulose) ⁶³, the stepwise reduction of methanol to CO_2 ⁶⁴, and the conversion of alcohols into amines ³², cellulosome scaffolds, like the dockerin-cohesin pair, have been designed.

Other protein domains, such as affibodies ⁶⁵, leucine zippers ⁶⁶, PDZ and SH3 domains ^{53,67}, CipB scaffolds ⁶⁷, and SpyTag-SpyCatcher system ⁴³ have also been successfully used to endow enzymes with spatial organization.

The SpyTag-SpyCatcher domain system created by the Howarth et al. is widely utilized for the co-localization of enzymes, and is one of the most illustrative instances of this protein scaffold (Figure 1.6b) ^{42,44,68}. SpyTag is a 13 amino acid long, short, unfolded peptide sequence made up of reactive aspartic acid residues that, when they recognize reactive lysine residues from their companion protein, SpyCatcher, produce covalent isopeptide bonds ⁶⁹. The SpyTag-SpyCatcher bioconjugation process's chemistry has been demonstrated to be extremely quick, highly effective, independent of its position on the protein sequence, and extremely robust in nature, with stability at a wide range of reaction conditions, including temperature (4-37 °C), pH (4-8), and the presence of various detergents ⁴². Therefore, the SpyTag-SpyCatcher system is an irreversible recombinant protein conjugation technique.

Due to its selectivity, irreversible covalent binding, and simplicity of use, the SpyTag-SpyCatcher conjugation system has a wide range of applications, such as the creation of covalently stabilized multi-protein complexes and enzyme cyclization ^{44,70}. This SpyTag-SpyCatcher protein system has also been used in protein scaffold mode for enzyme

targeting by engineering the bacterial microcompartment coat protein EutM of *Salmonella enterica*. A reversible system is obtained; therefore, it is easy to adapt and genetically program. This makes it a suitable system for simple self-assembly and subsequent immobilization of biocatalytic cascades⁴³.

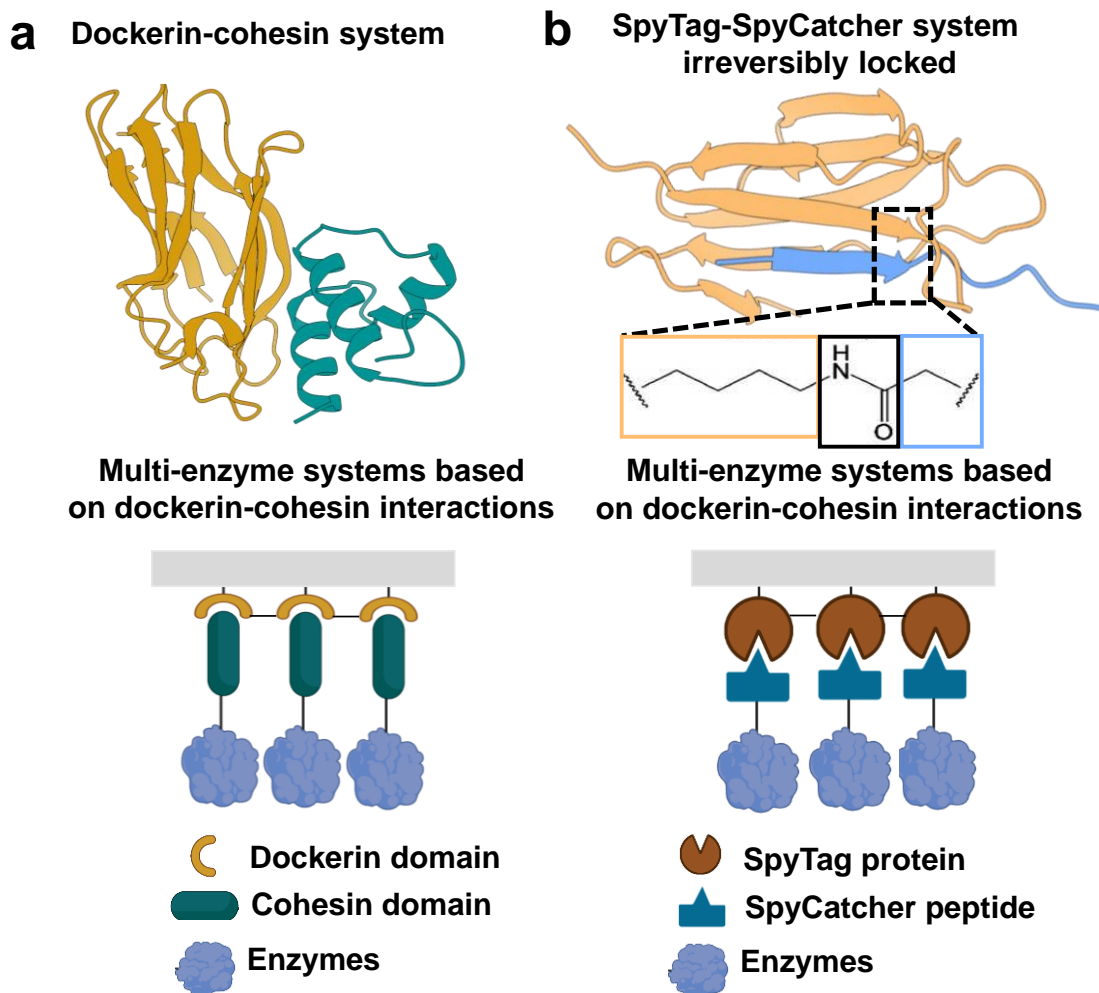


Figure 1.6. 3D structures and schematic representation of the most used protein-based scaffolds: Dockerin-cohesin and SpyTag-SpyCatcher systems. a) 3D structures and schematic representation of dockerin-cohesin pair (PDB ID: 1OHZ). The dockerin (in yellow) recognizes cohesin (in green) forming protein-protein interactions and gives rise to organized multi-enzyme systems upon immobilization on a solid support⁷¹. b) 3D structures, chemistry, and schematic representation of the SpyTag-SpyCatcher system (PDB ID: 4MLI). The SpyTag (in brown) recognizes its partner SpyCatcher (in blue) forming protein-peptide interactions and giving rise to organized multi-enzyme systems upon immobilization on solid support. The SpyTag-SpyCatcher system was formed in the presence of an irreversible isopeptide bond (black) between these two⁶⁹.

Consequently, the sequence of the whole construct determines each enzyme's spatial location, making these systems genetically tuneable. Comparing these co-immobilized enzymes on protein-based scaffolds to their non-co-immobilized counterparts, in which the enzymes are diluted in the reaction medium, has shown to be a great way to boost the catalytic efficiency of cascade reactions both within and outside of cells^{3,72}. Taking these protein-

based scaffolds as a source of inspiration, in this thesis we intend to use repeat proteins as synthetic protein-based scaffolds to co-immobilize enzymes in a precise and controlled spatial organization.

1.5. Repeat proteins as ideal scaffolds for enzyme immobilization

Repeat proteins are non-globular structures that participate in critical physiological processes and frequently function as scaffolds to promote protein-protein interactions. Repeat proteins are formed of a variable number of tandem repeats of a fundamental structural motif of 18 to 47 amino acids and are dominated by short-range, regular interactions^{73,74}. Thus, proteins create self-assembled complex nanostructures with a diversity of characteristics and activities^{74,75}. When the repeat protein structure is modular, each repeat unit can be used as a building block with uniquely defined attributes (stability, function, and inter-module interactions) to construct customized proteins and higher-order assemblies^{74,76}. Repeated proteins provide simplified systems where the relationships between sequence, structure, and function are well-understood. Consequently, by introducing changes to the sequence, these proteins enable the definition of the other two variables.⁷⁴ Moreover, these natural scaffolding systems have a large, exposed surface area and varied ratios in repeat treatment for scaffolding.

There are several families of repeat proteins made of units with various secondary structures, such as alpha-helical, beta-helical, or a mixture of both. Examples of repeat protein families are: Tetratricopeptide repeats (TPR), which consist of a 34 amino acid sequence that folds into a helix-turn-helix motif⁷⁷; ankyrin repeats (ANK), consisting of a 33 amino acid sequence that folds into a helix-loop-helix motif⁷⁸; and leucine-rich repeats (LRR), that consists of 20-30 amino acids that fold into a beta-turn-helix-turn-helix motif⁷⁹.

In this thesis, our focus has been on TPR proteins due to the extensive possibilities they offer in protein construction and engineering⁷⁷. The Regan Laboratory developed a consensus TPR sequence, known as **C**onsensus **T**etratricopeptide **R**epeat (CTPR), through statistical analysis of these natural proteins. This consensus sequence serves as a foundation for creating novel proteins while maintaining the crucial sequence-structure connection of the TPR fold⁸⁰ (Figure 1.7).

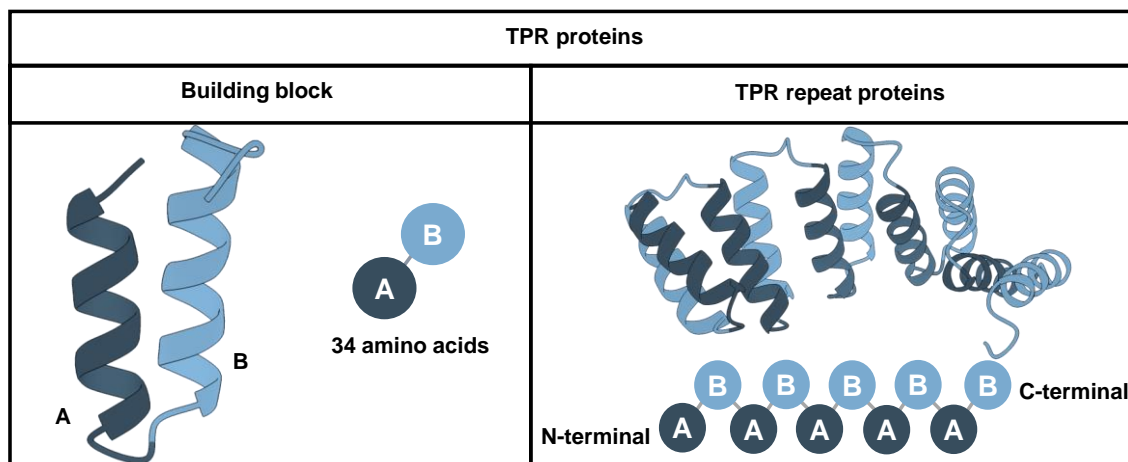


Figure 1.7. Schematic representation of TPR protein scaffolds. Left: 3D structure representation of the TPR building block conforming by A-B alpha helices. Right: The

crystal structure (PDB ID: 2AVP) of TPR protein by three repeats of each TPR building block are represented using the same schematic code as in the left panel.

CTPR proteins are a form of synthetic repeat protein that have remarkable features such as stability and toughness. Furthermore, since they are thermodynamically more stable than their natural counterparts, TPRs, they are more resistant to the destabilizing effects of mutations^{81,82}. Consequently, they are an excellent option for use as a protein scaffold. CTPRs can be repeated in tandem or connected to other peptides or proteins, such as enzymes, to encode larger CTPR proteins with super helical structures. The CTPR unit is a helix-turn-helix motif composed of 34 amino acids, of which only eight of them are required to ensure proper folding. Given that there are few conserved residues, adapting functional mutations to deliver the necessary properties is easier and more repeatable (Figure 1.8).

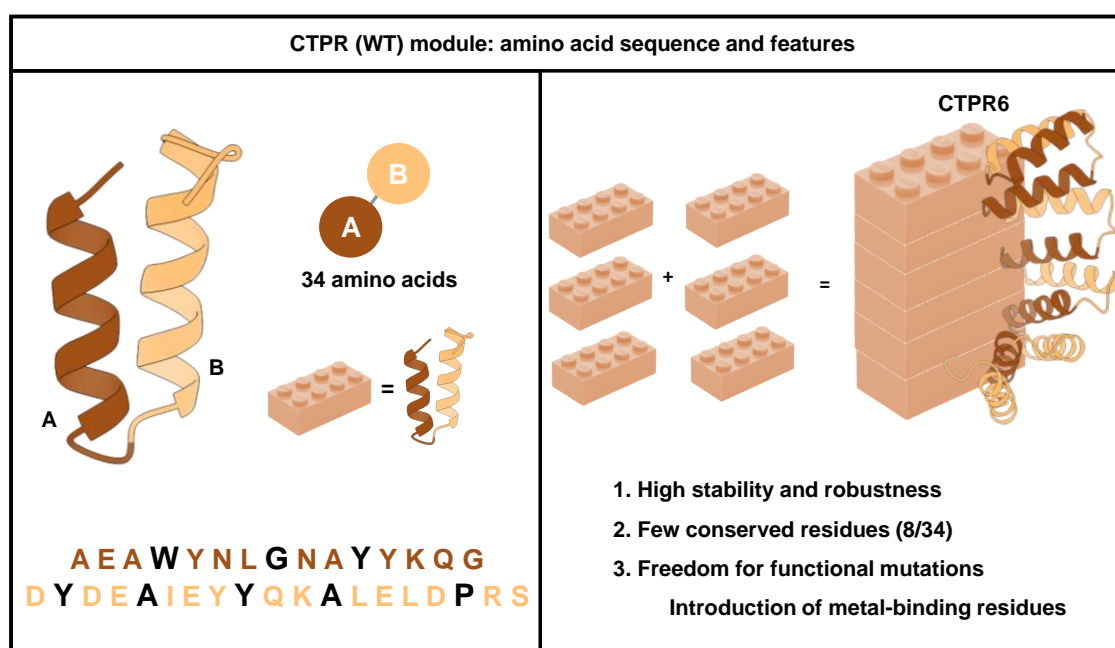


Figure 1.8. CTPR module: amino acid sequence and features. Left: 3D structure representation of the CTPR1 module consisting of 34 amino acids. The eight conserved residues are shown in black. Right: CTPR protein composed of 6 identical CTPR repeated modules. The main features related to CTPR proteins are high stability and robustness, few conserved residues (8/34), and freedom for functional mutations without compromising the protein structure, allowing for the introduction of metal-binding residues as an example.

In this thesis, CTPR proteins were chosen as scaffolds for enzyme immobilization because of their modular nature, which allows modifications at the repeat level. Not only do they build proteins with extensible structures, but they also have ample surface area to template molecules such as photoactive molecules⁸³, single-walled carbon nanotubes⁸⁴, electroactive clusters⁸⁵, and gold nanoparticles⁸⁶. In addition, as mentioned above only a few residues are conserved, allowing the remaining to be changed, allowing the encoding of desired functions.

The CTPR proteins tend to self-assemble through head-to-tail interactions in solution, crystal, and solid films⁸⁷. In addition, the introduction of single cysteines at the N- and C-terminals of the units allows the formation of disulfide bonds that staple these interactions, resulting in covalently linked linear nanofibers that cannot be spontaneously separated in aqueous solution^{87,88} (Figure 1.9). Taking this knowledge as a starting point, in this thesis we hypothesized the assembly of multi-enzyme systems using CTPR proteins as scaffolds.

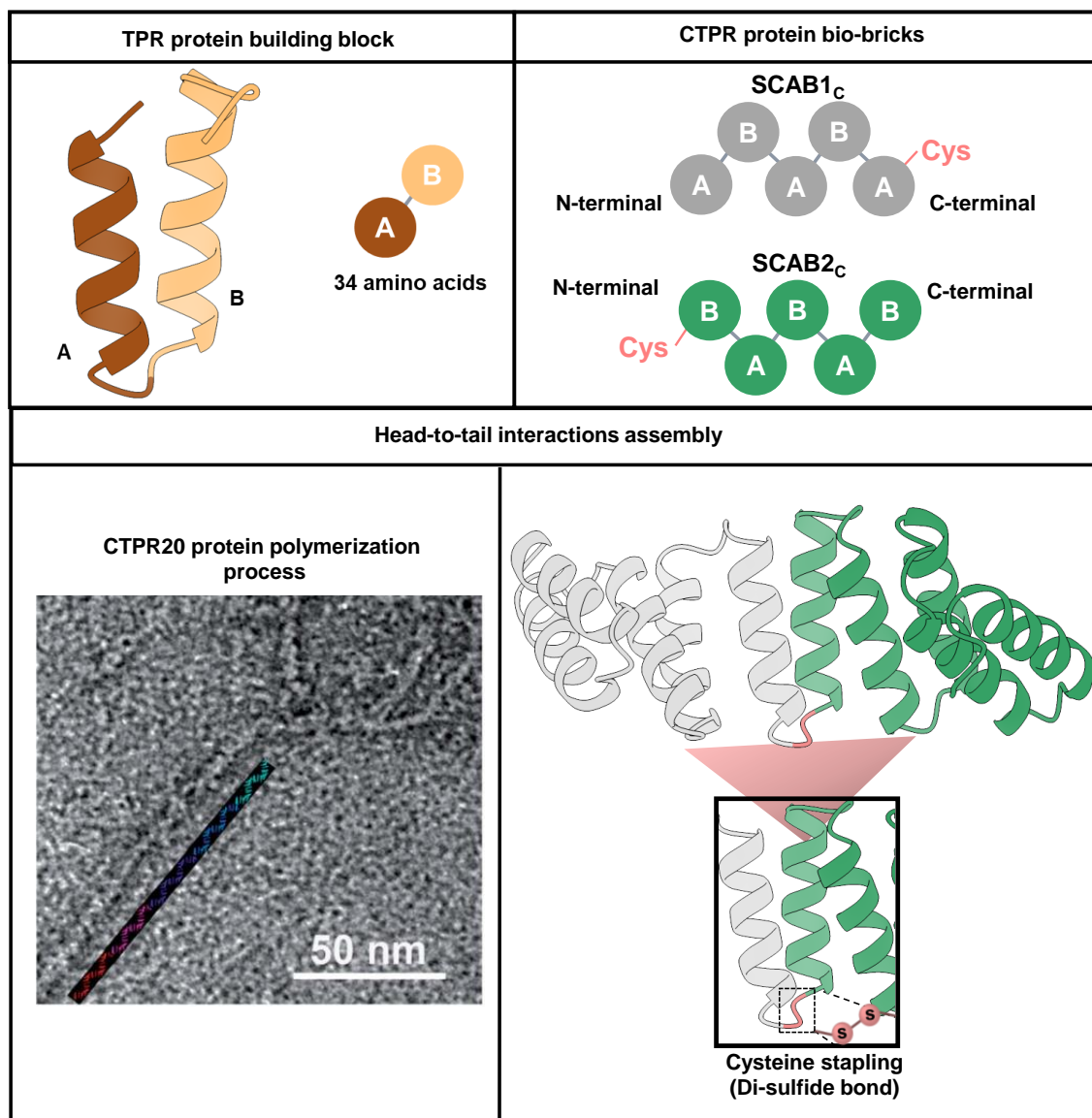


Figure 1.9. Schematic representation of head-to-tail interactions of CTPR proteins. Top left: 3D structure representation of the consensus sequence (CTPR). Top, right: Schematic representation of two CTPR proteins bio-bricks through helices A and B as follows: A-B-A-B-A (in grey) and B-A-B-A-B (in green). A single cysteine (Cys; in red) was introduced to each of the bio-bricks at the N- and C- terminus, respectively (Cys-A-B-A-B-A and Cys-B-A-B-A-B). Bottom left: CTPR20 linear polymeric structures. The schematic arrangement of CTPR20 units in the cladding fibers is shown with each CTPR20 unit displayed in different colors at approximately the scale of the fiber in the TEM image⁸⁷. Bottom, right: 3D structure representation two CTPR bio-bricks assembly by covalent

cysteine stapling promoted by the formation of a di-sulfide bond. This type of assembly is defined as assembly based on head-to-tail interactions.

The assemblies based on the coordination between metal-binding residues, such as histidines, cysteines, or even tyrosines with transition metals like copper, nickel, cobalt, or zinc, is another method that is gaining popularity to promote protein-protein assembly ⁷⁴. This method results in assemblies with different geometric structures based on the coordination between the amino acids that play the role of coordinators and the metals ^{89,90}.

Therefore, another way to assemble the enzymes on CTPR scaffolds is through the use of metal-driven assembly ⁸⁹⁻⁹² which is based on the interactions between designed coordination histidines within the scaffold and a metal ion (such as copper, nickel, or cobalt), resulting in the formation of a dimer that is stable in solution.

Finally, as an alternative, other repeat proteins, **Tetratricopeptide Repeat Affinity Proteins (TRAP)**, were as well used as scaffolds to assemble enzymes. TRAPs have conserved the design of their natural counterparts, TPRs. The TPR module binds its cognate peptide, MEEVD. The variations are in the binding domain to create three new TRAP modules (TRAP1, TRAP2, and TRAP3) which are going to bind three new cognate peptides: MEEVV, MERVW, and MRRWV. TRAP domains bind a selection of peptides with little cross-reactivity with other cellular proteins ⁹³. These TRAP domains were combined to create a single protein with up to three binding sites organized in sequence and the ability to orthogonally bind distinct enzymes tagged with their respective peptides confirmed by a 5 amino acid sequence. TRAP-based scaffolds are simple to modify by interface mutations, resulting in increased interaction flexibility ⁹³. The fusing of the short tag peptides to the enzymes (tagged enzymes) will allow the enzymes to be assembled in an orderly fashion on the TRAP-based scaffolds ^{93,94} (Figure 1.10).

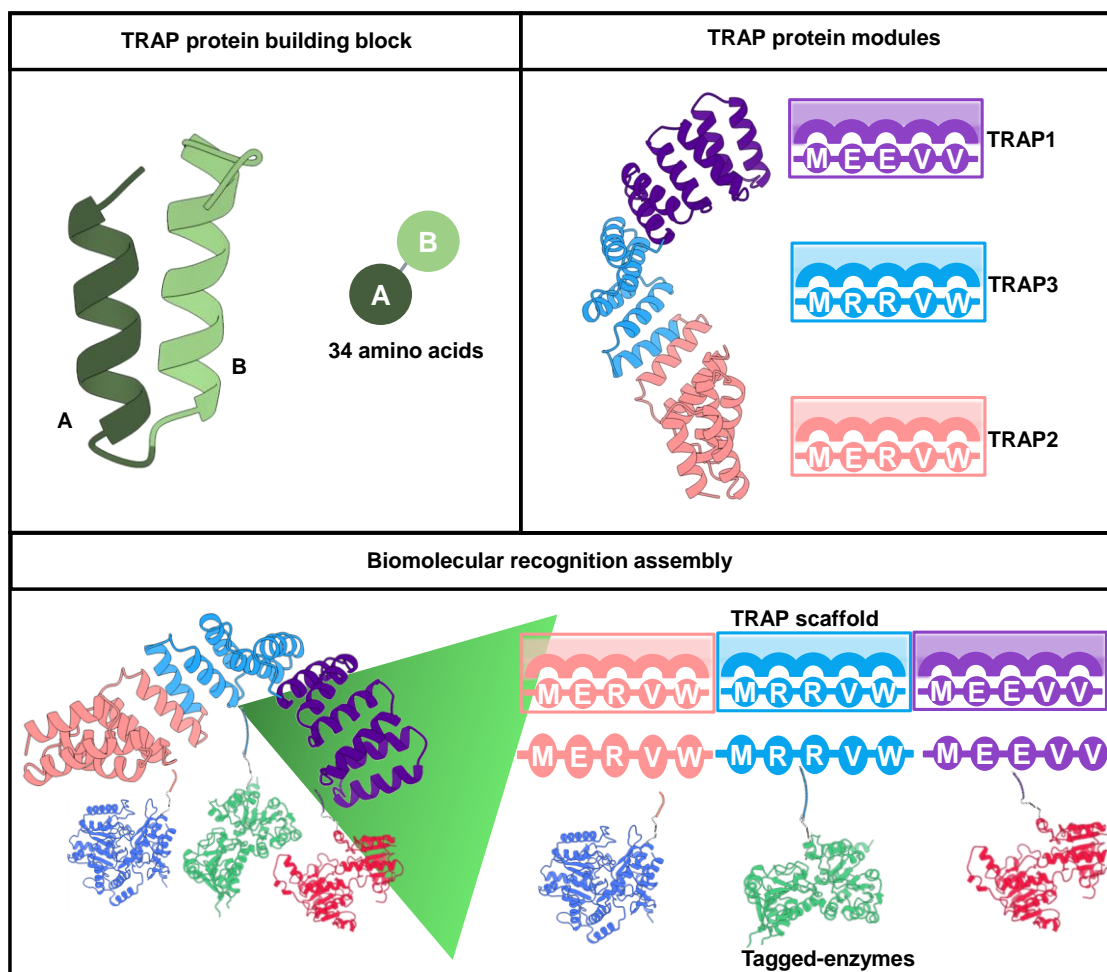


Figure 1.10. Schematic representation of TRAP proteins assembly through biomolecular recognition. Top left: 3D structure representation of the engineered sequence (TRAP). Top, right: 3D structure and schematic representation of the TRAP modules (TRAP1, TRAP2, and TRAP3; in purple, blue, and pink) and their affinity 5 amino acid peptides (MEEVV, MERVW, and MRRVW; in purple, blue and pink). Bottom: 3D structure and schematic representation of the enzyme assembly of TRAP modules and affinity peptides by biomolecular recognition.

As a result, TPR protein scaffolds can be an attractive and novel platform for enzyme assembly tools. These systems can encourage the development of long-term routes for the chemical production of high-added-value compounds.

Chapter 2. General aim and objectives

This thesis aims to contribute to the field of biocatalysis by designing synthetic proteins as scaffold units for the assembly of multi-enzyme pathways. Previous research has already demonstrated the advantages of immobilizing enzymes on biomacromolecule-based supports, such as proteins, including controlled immobilization and specific spatial organization. These factors have been shown to enhance catalytic activity and favor the substrate channeling effect.

Based on these findings, this thesis aims to further advance this field by integrating protein engineering tools to facilitate and precisely control enzyme immobilization. The work will focus on exploring the utilization of tetratricopeptide repeat (TPR) proteins, in particular, the consensus tetratricopeptide repeat (CTPR) module^{74,77,82} and tetratricopeptide repeat affinity proteins (TRAPs)⁹³, for the design of various scaffolds to assemble multi-enzyme pathways^{93,94}.

The research will encompass several scaffolding strategies, including the supramolecular assembly of CTPR proteins as scaffold building blocks, as well as the use of TRAPs designed for biomolecular recognition of orthogonal protein-peptide pairs.

It is hypothesized that the inclusion of synthetic TPR-protein scaffolds will not only enable the nanoscale arrangement of multi-enzyme systems but will also promote the spatial organization of functional modules, resulting in enhanced biocatalytic cascade reactions. By having the enzymes at nanometer distances, the presence of the scaffold is expected to facilitate substrate transport, resulting in more efficient catalytic reactions and enhanced substrate channeling effect^{10,13} (Figure 2.1). Furthermore, the TPR scaffolds possess binding sites that selectively and reversibly sequester reaction intermediates, such as cofactors, through electrostatic interactions. This sequestration increases the local concentration of intermediates and subsequently enhances catalytic efficiency.

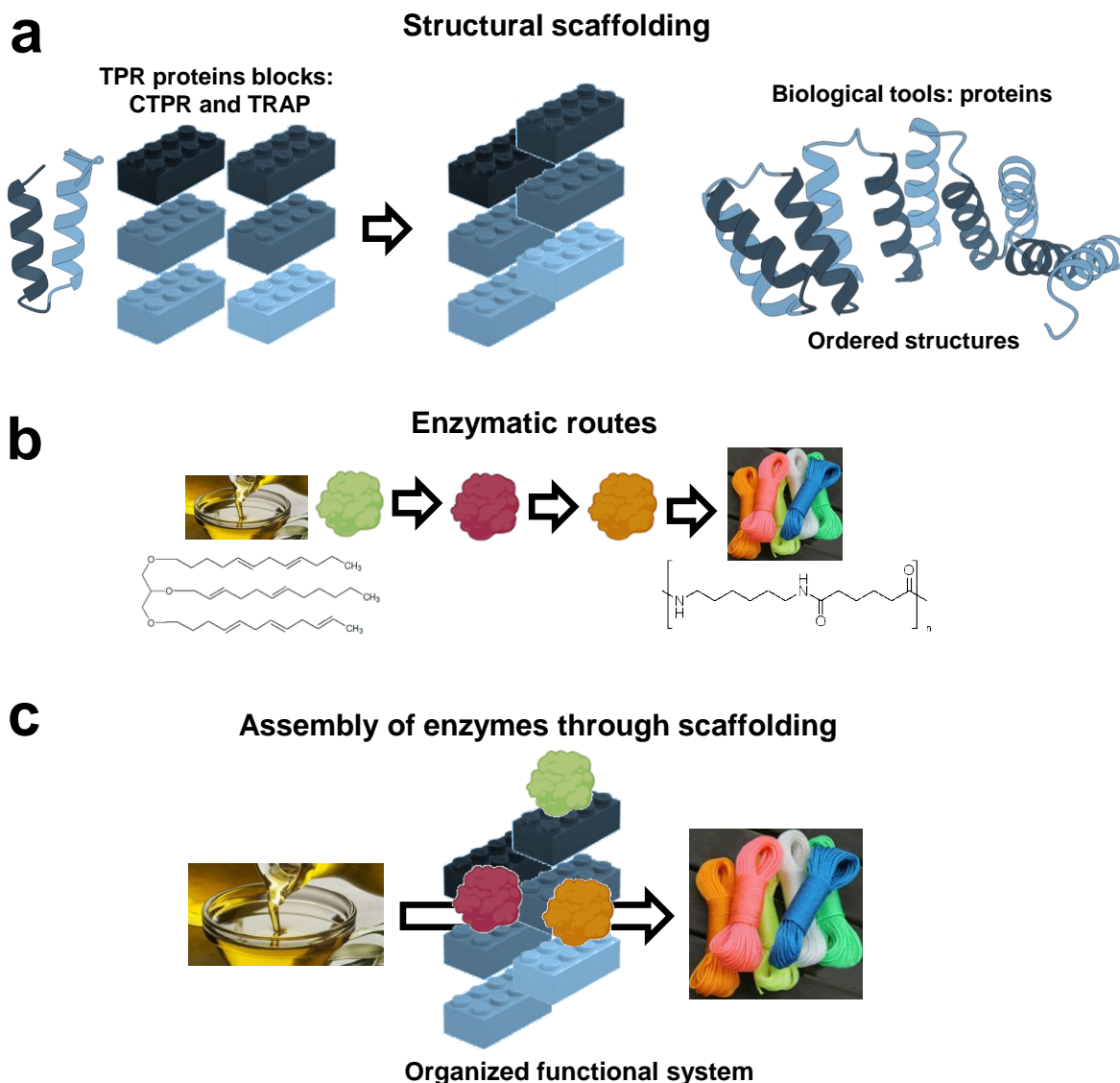


Figure 2.1. Nanometric arrangement of multi-enzyme systems. a) TPR protein blocks (CTPR and TRAP) as biological tools to obtain ordered structures. b) Enzyme engineering and enzymatic pathways. These kinds of systems result in long and multi-step processes. c) The assembly of enzymes through scaffolding originates an organized functional multi-enzyme system, that allows the confinement of several enzymes, with distinct functions, simplifying and reducing the number of steps for a given process.

Within these ordered multi-enzyme systems, two enzymatic cascades composed of two and three enzymes, respectively, will be constructed. The enzymes selected as model biocatalysts are formate dehydrogenase from *Candida boidinii* (FDH)⁹⁵, L-Alanine dehydrogenase from *Bacillus stearothermophilus* (AlaDH)⁹⁶, and ω -transaminase from *Pseudomonas fluorescens* (ω TA)⁹⁷. In the first multi-enzymatic cascade, the co-immobilized enzymes FDH and AlaDH will be used as biocatalysts. These two enzymes are known to simultaneously perform the asymmetric reduction of α -ketoacids to L-amino acids while recycling NADH using formate as an auxiliary electron donor and CO₂, as a by-product. In the second multi-enzymatic cascade, the co-immobilized enzymes FDH, ω TA, and AlaDH will be used as biocatalysts. These three enzymes carry out the

amination of benzaldehyde with *in-situ* recycling of electron (NADH) and amine (L-Alanine) donors.

Therefore, the main objective of this Ph.D. thesis is to develop synthetic proteins as scaffolding components for the assembly of multi-enzymatic pathways.

To achieve this objective, we will address the following specific objectives:

- **Develop two distinct orthogonal approaches for enzymatic scaffolding utilizing TPR proteins.** The first approach will focus on the supramolecular assembly of modified CTPR proteins, the newly engineered SCAffolding Bio-bricks (SCABs) will be developed based on CTPR units. The second approach will leverage biomolecular recognition assembly and will be based on Tetratricopeptide Repeat Affinity Proteins (TRAPs).
- **Validate the assembly of TPR modules.** For CTPR modules, orthogonal assembly strategies will be developed using techniques such as cysteine stapling and metal-directed assembly. The assembly process for TRAPs will rely on biomolecular recognition between enzyme-fused peptides and the corresponding orthogonal recognition sites within the TRAP scaffold.
- **Demonstrate the assembly of enzymes into TPR scaffolds (engineered CTPR and TRAP).** Engineered CTPR modules will be utilized to directly fuse specific dehydrogenase enzymes (FDH and AlaDH) to facilitate their assembly via supramolecular interactions. For biomolecular recognition-driven assembly, the selected dehydrogenase enzymes (FDH and AlaDH) will be fused to TRAP-tag peptides for consequent assembly on multivalent TRAP modules.
- **To investigate the impact of the assemblies on catalytic activity in organized multi-enzyme systems.** Firstly, we will verify the conservation of enzymatic activity in novel organized multi-enzyme systems. Secondly, we will examine how the integration of engineered TPR scaffolds enhances the development of highly efficient multi-enzymatic pathways in two industrially relevant biotransformations.

The objectives described above were addressed in the two experimental chapters (Chapter 4 and Chapter 5). The four specific objectives were successfully demonstrated in these chapters. While Chapter 4 focused on the complete development of ordered multi-enzyme systems using CTPR proteins as synthetic protein scaffolds, Chapter 5 achieved these objectives using TRAP proteins as synthetic protein scaffolds.

Chapter 3. Experimental procedures

The experimental procedures used in the experimental chapters of this thesis are described in detail below:

3.1. Protein Design, Cloning and Molecular Biology

3.1.1. SCAB_C modules

For the assembly of the SCAB_C modules, the wild-type (WT) CTPR3 (CTPR protein with 3 repetitions) gene cloned into the pProEx-HTa vector was used as a starting point. This DNA encodes a CTPR with three identical repeat modules (Table 3.1).

Table 3.1. The amino acid sequence of CTPR3 WT protein. CTPR3 WT amino acid sequence corresponding to A and B helices in dark and light brown respectively. The solvating helix of the CTPR3 WT protein in gray. The amino acid sequence corresponds to the hexa histidine-tag and the TEV protease cleavage site in black.

Protein	Amino acid sequence of the ORF
CTPR3-WT	MSYYHHHHHHDYDIPTTENLYFQGAM_DPGGNS AEAWYNLGNAYYKQG DYDEAIEYYQKALELDPNN AEAWYNLGNAYYKQG DYDEAIEYYQKALELDPNN AEAWYNLGNAYYKQG DYDEAIEYYQKALELDPNN AEAQNLGNAAKQKQG

A protocol based on the overlapping PCR technique was used to incorporate a coordinating cysteine at the C-terminal end into a CTPR3-WT gene and SCAB1_C was created. The same process was used for the creation of SCAB2_C. The only difference is that the coordinating cysteine was introduced at the N-terminal end. Then, the amplified fragments encoding the SCAB_C modules were cloned into the pProEx-HTa vector using the BamHI and HindIII restriction sites. Finally, the cysteine insertion in SCAB1_C and SCAB2_C was verified by sequencing (Table 3.2 and Figure 3.1).

Table 3.2. The amino acid sequence of SCAB1_C and SCAB2_C proteins. SCAB1_C and SCAB2_C amino acid sequences in grey and in green respectively. The cysteines introduced by overlapping PCR technique in red. The amino acid sequence corresponds to the Hexa histidine-tag and the TEV protease cleavage site in black.

Protein	Amino acid sequence of the ORFs
SCAB1_C	MSYYHHHHHHDYDIPTTENLYFQGAM_DPGGNS AEAWYNLGNAYYKQG DYDEAIEYYQKALELDPNN AEAWYNLGNAYYKQG DYDEAIEYYQKALELDPNN AEAWYNLGNAYYKQ G C
SCAB2_C	MSYYHHHHHHDYDIPTTENLYFQGAM_DPGGNS C DYDEAIEYYQKALELDPNN AEAWYNLGNAYYKQ GDYDEAIEYYQKALELDPNN AEAWYNLGNAYYKQ GDYDEAIEYYQKALELDPNN

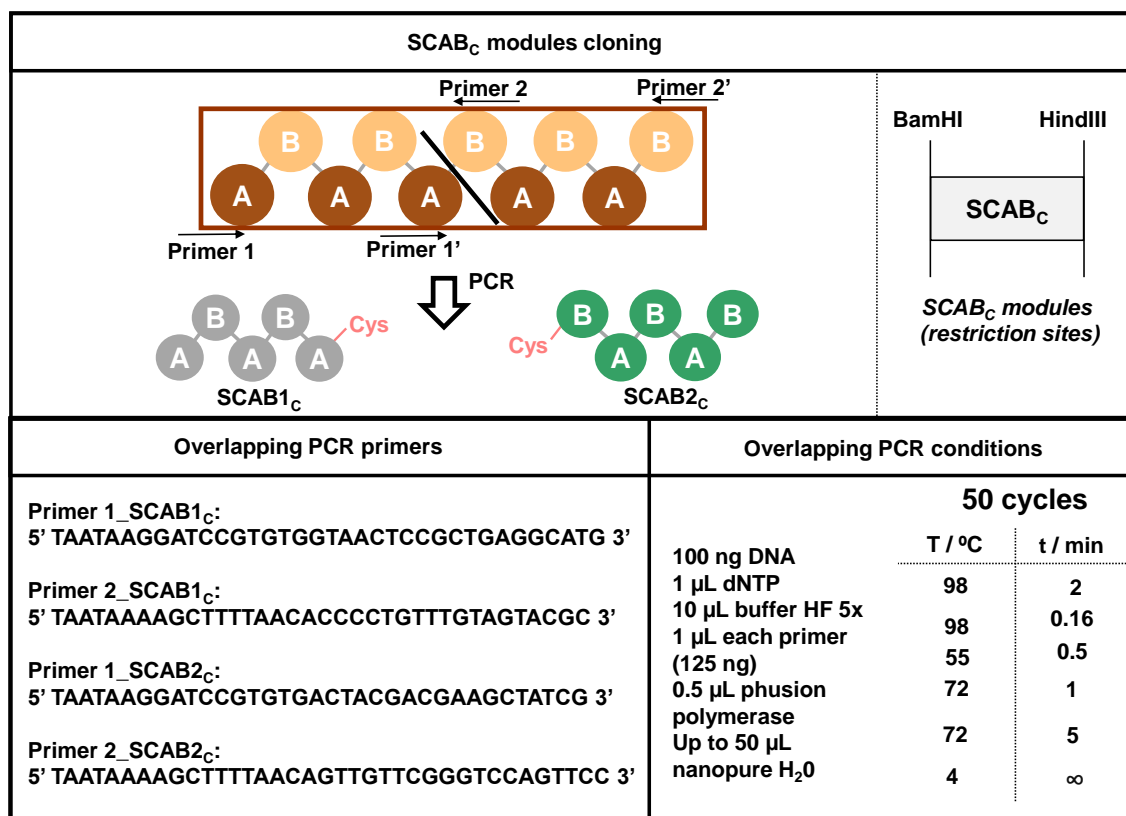


Figure 3.1. Schematic representation of the cloning process used for obtaining SCAB_C modules to perform assemblies through interactions of di-sulfide bonds by cysteine stapling. Starting from a CTPR3 WT, the cysteines at C- and N-terminal ends were introduced, thus obtaining the so-called SCAB_C modules, SCAB1_C and SCAB2_C modules (formed by 2 and a half CTPR repeats to seek orthogonality between the proteins for subsequent assembly by anchor cysteines). The restriction sites of the SCAB_C modules, the primers, and the conditions used in the overlapping PCR to perform the cloning process were shown.

3.1.2. SCAB_H modules

For the assembly of the SCAB_H modules, single point mutations were made in the CTPR1-WT gene (Table 3.3) cloned in the pProEx-HTa vector.

Table 3.3. The amino acid sequence of CTPR1 WT protein. CTPR1 WT amino acid sequence corresponding to A and B helices in dark and light brown respectively. The solvating helix of the CTPR1 WT protein in gray. The amino acid sequence corresponds to the hexa histidine-tag and the TEV protease cleavage site in black.

Protein	Amino acid sequence of the ORF
CTPR1-WT	MSYHHHHHHHDYDIPTTENLYFQGAMGS A EAWYNLGNAYYKQGDYDEAIEYYQKALELDRS AEAQN LGNAKQKQG

A sequential protocol based on rapid site-directed mutagenesis was used to incorporate four coordinating histidines into the CTPR1-WT (CTPR wild-type protein with 1 repetition)

gene and created CTPR1-4His. Within the 34 amino acids that the CTPR1-WT module is comprised of, mutations were introduced that included the substitution of CTPR1 WT residues for the coordinating histidines at positions E2H, N6H, N9H, and K13H. Three rapid targeted mutagenesis protocols⁹⁸ were performed to introduce the required mutations. In the first, the E2H mutation was introduced. In the second, the N6H and N9H mutations were introduced. In the latter, the K13H mutation was introduced. Then, amplified fragment encoding CTPR1-4His (Table 3.4) was cloned into pProEx-HTa vector using BamHI and HindIII restriction sites. Finally, after the cloning process, the colonies were sequenced to verify that the mutations had been introduced correctly. The histidines have been introduced to assemble two proteins by means of coordination between histidines and metal ions.

Table 3.4. The amino acid sequence of CTPR1-4His protein. CTPR1-4His amino acid sequence corresponds to A and B helices in dark and light brown respectively. The histidines introduced in A helix by quick change technique in cyan. The amino acid sequence corresponds to the Hexa histidine-tag and the TEV protease cleavage site in black.

Protein	Amino acid sequence of the ORF
CTPR1-4His	MSYYHHHHHHHDYDIPTTENLYFQGAMGS A H A W Y H L G H A Y L H Q G D Y D E A I E Y Y Q K A L E L D P R S AEA KQNLGN AKQKQG

After verifying the mutations in CTPR1-4His module, a modular cloning strategy, based on the restriction enzymes BamHI and BglII that generate compatible sticky ends⁹³, was used to construct two modules (SCAB1_H and SCAB2_H) with three repeats by combining the mutated CTPR1-4His with the coordination residues and the unmutated CTPR2 WT unit. SCAB1_H gene was created by ligation of the mutated fragment of CTPR1-4His with the coordination histidines into the pProEx-HTa vector, which already contained the wild-type fragment of a CTPR2 WT gene only digested with BamHI. SCAB2_H gene was created by ligation of a wild-type of fragment of a CTPR2 WT into the pProEx-HTa vector, which already contained the mutated CTPR1-4His gene with the coordination histidines only digested with BamHI. The insertion of the histidine-mutated fragment in SCAB1_H and SCAB2_H was verified by sequencing (Table 3.5 and Figure 3.2).

Table 3.5. The amino acid sequence of SCAB1_H and SCAB2_H proteins. SCAB1_H and SCAB2_H amino acid sequences in dark pink and in respectively. The histidine modules introduced by quick change technique in cyan. The amino acid sequence corresponds to the Hexa histidine-tag and the TEV protease cleavage site in black.

Protein	Amino acid sequence of the ORFs
SCAB1_H	MSYYHHHHHHHDYDIPTTENLYFQGAMGS A E A W Y N L G N A Y Y K Q G D Y D E A I E Y Y Q K A L E L D P R S A E A W Y N L G N A Y Y K Q G D Y D E A I E Y Y Q K A L E L D P R S A H A W Y H L G H A Y L H Q G D Y D E A I E Y Y Q K A L E L D P R S
SCAB2_H	MSYYHHHHHHHDYDIPTTENLYFQGAMGS A H A W Y H L G H A Y L H Q G D Y D E A I E Y Y Q K A L E L D P R S A E A W Y N L G N A Y Y K Q G D Y D E A I E Y Y Q K A L E L D P R S A E A W Y N L G N A Y Y K Q G D Y D E A I E Y Y Q K A L E L D P R S

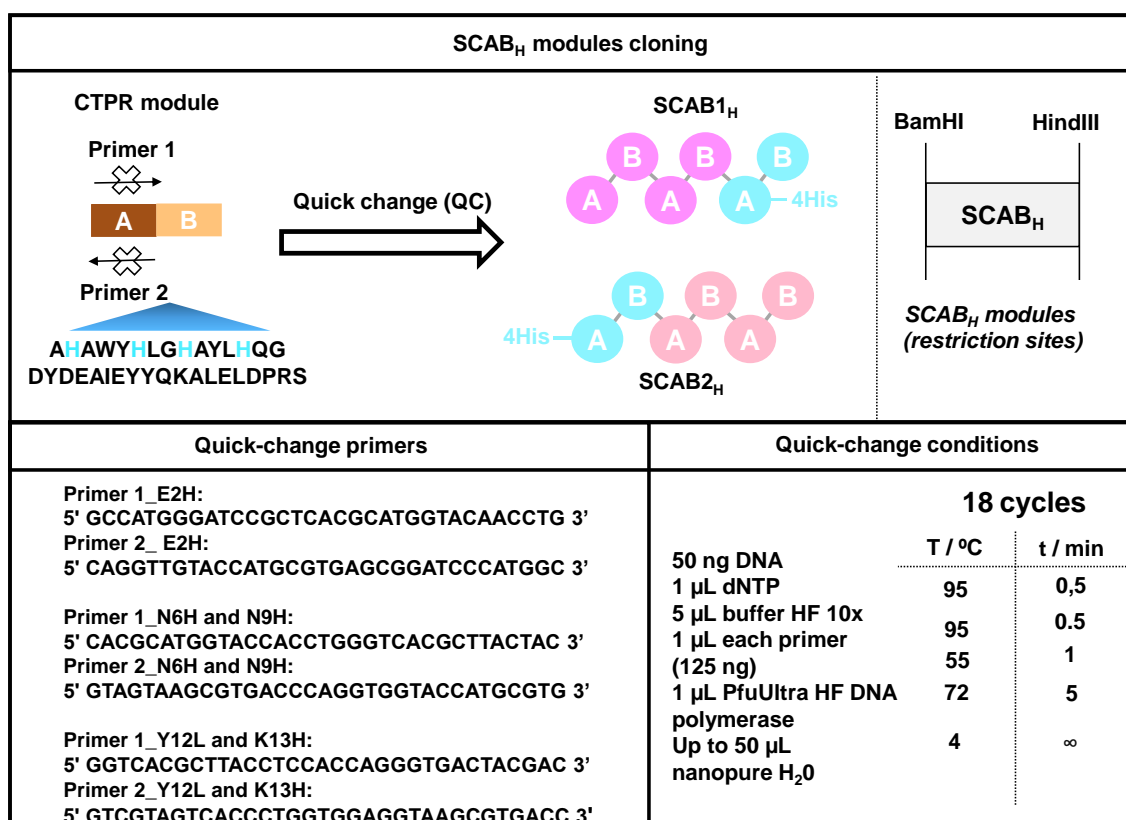


Figure 3.2. Schematic representation of the cloning process used for obtaining SCAB_H modules to perform assemblies through interactions between coordination histidines and metals. Starting from a CTPR1-WT, the 4 histidines are introduced to obtain a CTPR-4His module which is cloned by block cloning to a CTPR2-WT at the N- and C- terminal ends obtaining SCAB_{1H} and SCAB_{2H}, respectively. Thus, SCAB_H modules consisting of 3 CTPR repeats one of each with 4 histidines are obtained. The restriction sites of the SCAB_H- modules, the primers and the conditions used in the quick change (QC) to perform the cloning process were shown.

3.1.3. SCAB-enzyme fusions

Bioassays, an industrial partner of the HOMBIOCAT project, fused the SCAB_C and SCAB_H modules to FDH and AlaDH enzymes and cloned into pet-28b (+) plasmid using our previously constructed plasmids containing each part of these SCAB-enzyme fusions (FDH@SCAB_{1C}, AlaDH@SCAB_{2C}, FDH@SCAB_{1H} and AlaDH@SCAB_{2H}). To clone the enzymes into pet-28b (+) vector the fragment encoding the enzyme was digested with HindIII/BamHI to be subsequently inserted into the scaffold plasmid with the SCAB modules were digested with HindIII/BglII to make a fusion protein harboring the scaffold domain at the C-terminal end (Table 3.6).

Table 3.6. The amino acid sequences of the SCAB-enzyme fusions. SCAB_{1C} and SCAB_{2C} amino acid sequences in grey and in green respectively. The cysteines were introduced by overlapping PCR technique in red. FDH and AlaDH amino acid sequences in turquoise and in purple respectively. SCAB_{1H} and SCAB_{2H} amino acid sequences in dark pink and in respectively. The histidine modules introduced by quick change technique in cyan. FDH and AlaDH amino acid sequences in turquoise and in purple respectively.

The linker between the SCAB modules and the enzymes in yellow. The amino acid sequence corresponds to the Hexa histidine-tag and the TEV protease cleavage site in black.

Protein	Amino acid sequence of the ORFs
FDH@ SCAB1_c	<p>MGSSHHHHHHSSGENLYFQGHMGNS</p> <p>AEAWYNLGNAYYKQG DYDEAIEYYQKALELDPNN AEAWYNLGNAYYKQG DYDEAIEYYQKALELDPNN AEAWYNLGNAYYKQG CGSGSGSKLKRSMKIVLV LYDAGKHA ADEEKLYGCTENKLGIANWLKDQGH ELITTS DKEGETSEL DKHIPDADIIITTPFHPAYITK ERLDKAKNLKLVVAGV GSDHIDL DYINQTGKKIS VLEVTGSNVVSV A EHVVM TMLVLRNFVPAHEQII NHDWEVA AIAKDAYDIEG KTIATIGAGRIGYRVLE RLLPFNPKELLYDYQALPKEAEEKVGARRVENI EELVAQADIVTVNAPLHAGTKGLINKELLSKFKKG AWLVNTARGAICVAEDVAAALESGQLRGGYGDV WFPQPAPKDH PWRDMRNKYGAGNAMTPHYS GTT LDAQTRYAEGTKNILESFFT GKFDYRPQDIILLNG EYVTKAYGKHDKKK</p>
AlaDH@ SCAB2_c	<p>MGSSHHHHHHSSGENLYFQGHMGNS</p> <p>CDYDEAIEYYQKALELDPNNAEAWYNLGNAYYKQ GDYDEAIEYYQKALELDPNNAEAWYNLGNAYYKQ GCDYDEAIEYYQKALELDPNNGSGSGSKLKRSIIG VPKEIKNNENRVALTPGGVSQLISNGHRVLEVETG AGLGSGFENEAYESAGAEIIADPKQVWDAEMVMK VKEPLPEEYVYFRKGLVLF TYLHLAAEPELAQALK DKGVTAIAYETVSEGR TLP LTPMSEVAGRMAAQ IGA QFLEKPKGGK GILLAGVPGVSRGKV TIIGGGV VGTNAAKMAVGLGADVTIIDLNADRLRQLDDIFGH QIKTLISNPVNIADAVA EADLLICAVLIPGAKAPTL VTEEMVKQMKPGSVIDVAIDQGGIVETVDHITTH DQPTYEKHGVVHYAVANMPGAVPRTSTIALTNVT VPYALQIANKGAVKALADNTALRAGLNTANGHVT YEA VARDLGYEYVPAEKALQDESSVAGA</p>
FDH@ SCAB1_H	<p>MGSSHHHHHHSSGENLYFQGHMGNS</p> <p>KIVLVLYDAGKHA ADEEKLYGCTENKLGIANWLK DQGH ELITTS DKEGETSEL DKHIPDADIIITTPFHP AYITKERLDKAKNLKLVVAGV GSDHIDL DYINQT GKKISVLEVTGSNVVSV A HVVM TMLVLRNFVPA HEQIINH DWEVA AIAKDAYDIEG KTIATIGAGRIGY RVLERLLPFNPKELLYDYQALPKEAEEKVGARR VENIEELVAQADIVTVNAPLHAGTKGLINKELLSK FKKGAWLVNTARGAICVAEDVAAALESGQLRGGY GDVWFPQPAPKDH PWRDMRNKYGAGNAMTPHYS GTTLD A QTRYAEGTKNILESFFT GKFDYRPQDIIL LNGEYVTKAYGKHDKKK GSGSGSAEAWYNLGN AYYKQGDYDEAIEYYQKALELDPRSAEAWYNLGN AYYKQGDYDEAIEYYQKALELDPRSAHAWYHLGHA YLHQGDYDEAIEYYQKALELDPRS</p>
AlaDH@ SCAB2_H	<p>MSYYHHHHHHHDYDIPTTENLYFQGGAMGS</p> <p>AHAWYHLGHAYLHQGDYDEAIEYYQKALELDPRS AEAWYNLGNAYYKQGDYDEAIEYYQKALELDPRS AEAWYNLGNAYYKQGDYDEAIEYYQKALELDPRS GSGSGSIIGVPKEIKNNENRVALTPGGVSQLISNG HRVLEVETGAGLGSGFENEAYESAGAEIIADPKQVW</p>

	<p>DAEMVMKVKEPLPEEYVYFRKGLVLFITYLHLAAE PELAQALKDKGVTAIAYETVSEGRTLPLLLTPMSEV AGRMAAQIGAQFLEKPKGGKGILLAGVPGVSRGK VTIIGGGVVGTAAKMAVGLGADVTTIIDLNADRLR QLDDIFGHQIKTLISNPVNIADAVAEADLLICAVLIP GAKAPTLVTEEMVKQMKPGSVIVDVAIDQGGIVE TVDHITTHDQPTYEKHGVVHYAVANMPGAVPRTS TIALTNVTVPYALQIANKGAVKALADNTALRAGLN TANGHVTYEAVARDLGYEYVPAEKALQDESSVAGA</p>
--	--

3.1.4. TRAP scaffolds and tagged enzymes

TRAP1-3 and TRAP2-3-1 scaffolds are a combination of two or three Tetratricopeptide Repeat Affinity Proteins (TRAP). Each TRAP represents an engineered TPR module derived from the TPR2A domain of HOP protein⁹⁹ evolve to bind a cognate peptide MEEVV, MERVW, and MRRVW for TRAP1, TRAP2, and TRAP3, respectively⁹³ (Table 3.7). TRAP1-3 and TRAP2-3-1 genes (Table 3.7) were purchased from Biomatik, USA. These genes were cloned in the pet-28b (+) vector, which has been selected as the standard vector for the expression of the proteins. The genes were ordered flanked with four restriction sites for cloning in pet-28b (+) or in pProEx-HTa vector, a plasmid commonly used to express TPR proteins⁸¹.

Table 3.7. The amino acid sequences of the TRAP modules and TRAP scaffolds. TRAP1, TRAP2, and TRAP3 modules amino acid sequences in purple, in pink and in blue respectively. GGSSLQ is the sequence for the linker between the TRAP modules in yellow. The amino acid sequence corresponds to the Hexa histidine-tag and the Thrombine protease cleavage site in black.

Protein	Amino acid sequence of the ORFs
TRAP1	<p>MSYYHHHHHHHDYDIPTTENLYFQGSKQ ALKEKELGNDAYKKKDFDTALKHYDKAKELDPTN MYILNQA AVYFEKGDYNKCRELCEKAIEVGRENR EDYRLIAIAYARIGNSYFKEEKYKDAIHFNKSLA EHRTPKVLKKCQQA EKILKEQ</p>
TRAP2	<p>MSYYHHHHHHHDYDIPTTENLYFQGSKQ ALKEKELGNDAYKKKDFDTALKHYDKAKELDPTN MYIMNQA AVYFEKGDYNKCRELCEKAIEVGRE NREDYRMIA YAYARIGNSYFKEEKYKDAIHFNKSLA LAEHRTPKVLKKCQQA EKILKEQ</p>
TRAP3	<p>MSYYHHHHHHHDYDIPTTENLYFQGSKQ ALKEKELGNDAYKKKDFDTALKHYDKAKELDPTN MYIMNQA AVYFEKGDYNKCRELCEKAIEVGRENR EDYRMIA YAYADIGDSYFKEEKYKDAIHFNKSLA EHRTPKVLKKCQQA EKILKEQ</p>
TRAP1-3	<p>MSYYHHHHHHHDYDIPTTENLYFQGSKQ ALKEKELGNDAYKKKDFDTALKHYDKAKELDPTN MYILNQA AVYFEKGDYNKCRELCEKAIEVGRENR EDYRLIAIAYARIGNSYFKEEKYKDAIHFNKSLA EHRTPKVLKKCQQA EKILKEQ GGSSLQ ALKEKEL GNDAYKKKDFDTALKHYDKAKELDPTNMYIMNQA AVYFEKGDYNKCRELCEKAIEVGRENRREDYRMIA</p>

	<p style="text-align: center;">YAYADIGDSYFKEEKYKDAIHFNKSLAEHRTPKV LKKCQQA EKILKEQ</p>
TRAP2-3-1	<p style="text-align: center;">MSYYHHHHHHDYDIPTTENLYFQGSKQ</p> <p style="text-align: center;">ALKEKELGNDAYKKKDFDTALKHYDKAKELDPTN MYIMNQA AVYFEKGDYNKCRELCEKAIEVREN EDYRMIAYAYARIGNSYFKEEKYKDAIHFNKSLA EHRTPKVLKKCQQA EKILKEQGGSGLQALKEKEL GNDAYKKKDFDTALKHYDKAKELDPTNMYILNQA AVYFEKGDYNKCRELCEKAIEVRENREDYRLIAI AYARIGNSYFKEEKYKDAIHFNKSLAEHRTPKVL KKCQQA EKILKEQGGSGLQALKEKELGNDAYKKK DFDTALKHYDKAKELDPTNMYIMNQA AVYFEKGD YNKCRELCEKAIEVRENREDYRMIAYAYADIGD SYFKEEKYKDAIHFNKSLAEHRTP KVLKKCQQA EKILKEQ</p>

The enzymes FDH and AlaDH already cloned in pet-28b (+) vector were tagged with the peptides that recognize TRAP1 and TRAP3 modules (MEEVV in the case of FDH enzyme and MRRVW in the case of AlaDH enzyme) by overlapping PCR. The amplified fragments encoding the tagged enzymes were cloned into pet-28b (+) vector using NdeI and XhoI restriction enzymes. The fragment encoded ω TA enzyme was directly cloned into pet-28b (+) vector which contains the recognition peptide for TRAP2, MERVW using NdeI and HindIII restriction enzymes (Table 3.8 and Figure 3.3).

Table 3.8. The amino acid sequences of the TRAP tag peptides and tagged enzymes. FDH1, ω TA2 and AlaDH3 amino acid sequences in red, in blue and in green respectively. KLGSGSGSDDTSR is the sequence for the linker between the enzymes and TRAP tag peptides modules in yellow. The amino acid sequence corresponds to the Hexa histidine-tag and the TEV protease cleavage site in black.

Protein	Amino acid sequence of the ORFs
TRAP1 peptide	KLGSGSGSDDTSR MEEVV
TRAP2 peptide	KLGSGSGSDDTSR MERVW
TRAP3 peptide	KLGSGSGSDDTSR MRRVW
FDH1	<p style="text-align: center;">MGSSHHHHHSSGLVPRGSHMEFKRS</p> <p style="text-align: center;">MKIVLVLYDAGKHA ADEEKLYGCTENKLGIANWLK DQGHELITTS DKEGETSEL DKHIPDADIITTPFHP AYITKERLDKAKNLKLVVAVGSDHIDL D YINQTG KKISVLEVTGSNVVSVAEHVMTMLVLVRNFVPAH EQIINH DWEVA AIAKDAYDIEGKT IATIGAGRIGYR VLERLLPFNP KELLYYDYQALPKEAEEKVGARRVE NIEELVAQADIVTVNAPLHAGTKGLINKELLSKFKK GAWLVNTARGAICVAEDVAAALESGQLRGGYGGDV WFPQPAPKDHPWRDMRNKYGAGNAMTPHYS GTT LDAQTRYAEGTKNILESFFT GKFDYRPQDIILLNGE YVTKAYGKHDKKKKLGSGSGSDDTSR MEEVV</p>

<i>ωTA2</i>	<p>MGSSHHHHHHSSGENLYFQGHMLRS</p> <p>NSNNKAWLKEHNTVHMMHPMQDPKALHEQRPLII</p> <p>QSGKGVHITDVDGRRFIDCQGGLWCVNAGYGRRE</p> <p>IIDAVTRQMEELAYYSLFPGSTNAPAIALSQKLTEV</p> <p>AAEEGMVKASFGLGGSDAVETALKIARQYWKLEG</p> <p>QPDKVKFVSLYNGYHGLNFGGMSACGGNAWKSS</p> <p>YEPLMPGFFQVESPHLYRNPFTNDPEELAEICAQI</p> <p>LERQIEMQAPGTVAALIAEPIQGAGGVIVPPASYWP</p> <p>RLRQICDKYDILLIADEVITGLGRSGSLFGSRGWGV</p> <p>KPDIMCLAKGISSGYVPLSATLVNSRVARAWERDA</p> <p>GFTSVYMHGYTYSGHPVSCAAALAAIDIVLQENLA</p> <p>ENARVVDYFLEKLLILKDKHRAIGDVRGKGLMLA</p> <p>VELVKERATKEPFGPADAYPLAISEACVNNGVMIR</p> <p>TIVNKLIISPPLTFTTEHVDEIVLDRAFVANPWK</p> <p style="text-align: center;">LGSGSGSDDTSRMERVW</p>
<i>AlaDH3</i>	<p>MGSSHHHHHHSSGLVPRGSHMEFKRS</p> <p>IIGVPEIKNNENRVALTGGVSQLISNGHRVLVET</p> <p>GAGLGSGFENEAYESAGAEIIADPKQVWDAEMVM</p> <p>KVKEPLPEEYVYFRKGLVLFITYLHLAAEPELAQAL</p> <p>KDKGVTAIAYETVSEGRTPLLTPMSEVAGRMAAQ</p> <p>IGAQFLEKPKGGKILLAGVPGVSRGKVTIIGGGV</p> <p>VGTNAAKMAVGLGADVTIIDLNADRLRQLDDIFGH</p> <p>QIKTLISNPVNIADAVAEADLLICAVLIPGAKAPTLV</p> <p>TEEMVKQMKPGSVIVDVAIDQGGIVETVDHITTHD</p> <p>QPTYEKHGVVHYAVANMPGAVPRTSTIALTNVTVP</p> <p>YALQIANKGAVKALADNTALRAGLNTANGHVTYEA</p> <p>VARDLGYEYVPAEKALQDESSVAGAKLGS</p> <p style="text-align: center;">DTSRMRRVW</p>

Tagged enzymes cloning																									
Overlapping PCR primers	Overlapping PCR conditions																								
<p>Primer 1_T7 forward: 5' TAATACGACTCACTATAGG 3'</p> <p>Primer 2_FDHD: 5'CGGTAAACACGATAAAAAAAAAAAGCTTGGCTCTG GCAGCGGAAGCGATGATACGAGCCGCATGGAGGAA GTCGTGTGAGGATCCAAACTCGAG 3'</p> <p>Primer 2_LAlaDH: 5'GTGGTGCTCGAGTTTGGATCCTCACCAAACGCGGC GCATACGGCTGGTATCGTCCGAGCCACTCCCGCTTCC AAGTTCGCGCCAGCCACGGAC 3'</p>	<p>50 cycles</p> <table border="1"> <thead> <tr> <th></th> <th style="text-align: center;">T / °C</th> <th style="text-align: center;">t / min</th> </tr> </thead> <tbody> <tr> <td>100 ng DNA</td> <td style="text-align: center;">98</td> <td style="text-align: center;">2</td> </tr> <tr> <td>1 μL dNTP</td> <td style="text-align: center;">98</td> <td style="text-align: center;">0.16</td> </tr> <tr> <td>10 μL buffer HF 5x</td> <td style="text-align: center;">98</td> <td style="text-align: center;">0.5</td> </tr> <tr> <td>1 μL each primer (125 ng)</td> <td style="text-align: center;">55</td> <td style="text-align: center;">0.5</td> </tr> <tr> <td>0.5 μL phusion polymerase</td> <td style="text-align: center;">72</td> <td style="text-align: center;">1.30</td> </tr> <tr> <td>Up to 50 μL nanopure H₂O</td> <td style="text-align: center;">72</td> <td style="text-align: center;">5</td> </tr> <tr> <td></td> <td style="text-align: center;">4</td> <td style="text-align: center;">∞</td> </tr> </tbody> </table>		T / °C	t / min	100 ng DNA	98	2	1 μL dNTP	98	0.16	10 μL buffer HF 5x	98	0.5	1 μL each primer (125 ng)	55	0.5	0.5 μL phusion polymerase	72	1.30	Up to 50 μL nanopure H ₂ O	72	5		4	∞
	T / °C	t / min																							
100 ng DNA	98	2																							
1 μL dNTP	98	0.16																							
10 μL buffer HF 5x	98	0.5																							
1 μL each primer (125 ng)	55	0.5																							
0.5 μL phusion polymerase	72	1.30																							
Up to 50 μL nanopure H ₂ O	72	5																							
	4	∞																							

Figure 3.3. Schematic representation of the cloning process used for obtaining tagged enzymes to perform assemblies based on biomolecular recognition. A small peptide is directly fused by overlapping PCR at the C-terminal end based on enzymes giving rise to enzyme- and peptide-conformation tagged-enzymes. The restriction sites of the tagged enzymes, the primers and the conditions used in the PCR to perform the cloning process were shown.

3.2. Protein expression and purification

SCAB modules, SCAB-enzyme fusions, TRAP proteins and tagged enzymes (enzymes fused to peptides), were overexpressed in *Escherichia coli* C41 cells. An overnight saturated cell culture was diluted in 1 L of LB and grown to an OD₆₀₀ = 0.6-0.7 at 37 °C. The next step was to induce overexpression by the addition 0.6 mM Isopropyl β-d-1-thiogalactopyranoside (IPTG) followed by overnight growth at 20 °C. After overnight growth, the cells were centrifuged for 15 min at 4500 rpm and at a temperature of 4 °C. Pellets acquired after centrifugation of each liter were resuspended with 30 mL of TRIS-HCl buffer (500 mM NaCl, 5 mM imidazole, 50 mM TRIS-HCl pH 8.0) and frozen overnight at -20 °C. SCAB_C modules and SCAB_C-enzyme fusions in the resuspension buffer contained 5 mM final concentration of the reducing agent, Dithiothreitol (DTT). After overnight freeze, the resuspended pellet was defrosted and incubated for 30 min on ice with the addition of lysozyme (1 mg per mL of solution), 1 aliquot (50 μL) of protease inhibitor and 5 μL of DNase. After 30 min, cold sonication was performed for a total of 20 min with 2 cycles of 10 min, 40 % amplitude and 1 s on / 1 s off cycles. Between the two cycles the ice was changed to maintain the cold chain. The lysate was then centrifuged at 10000 rpm for 45 min at 4 °C. After centrifugation, the proteins were purified. All the proteins (SCAB modules, SCAB-enzyme fusions, TRAP proteins and tagged enzymes) were purified as His-tag fusions following standard protocols using nickel nitrilotriacetic acid affinity chromatography (Ni-NTA)¹⁰⁰. Before starting the protein purification with the Ni-NTA column, the column was equilibrated with TRIS-HCl lysis buffer (5 mM imidazole, 50 mM TRIS-HCl pH 8.0) to find the optimal conditions. First step was running supernatant through Ni-NTA column and collect (the protein sticks to the column due to the His-Tag); then washed the column with 10 volumes (50 mL) of TRIS-HCl wash buffer (800 mM NaCl, 20 mM imidazole, 50 mM TRIS-HCl pH 8.0) relative to the volume of the Ni-NTA column (5 mL) and collect. Finally, ran elution buffer 5 volume (25 mL) of TRIS-HCl elution buffer (500 mM NaCl, 300 mM imidazole, 50 mM TRIS-HCl pH 8.0), and collect: it was eluted our proteins. It should be noted that in the proteins of the purifications that have cysteines all buffers carried 5 mM final concentration of the reducing agent, DTT. The His-tag was removed using tobacco etch virus (TEV) protease for all the proteins except for SCAB_{1C}, FDH@SCAB_{1C}, TRAP1-3 and TRAP2-3-1 which maintained the His-Tag to immobilize the final assembly on solid supports.

Proteins that had His-Tag cleaved by TEV protease followed the following purification protocol. The eluted proteins were mixed with a final concentration of 2 mM EDTA, 5 mM DTT together with an aliquot of TEV protease (TEV:protein ratio 1:50) and incubated at 4 °C overnight. After overnight incubation, they were purified by Ni-NTA column except for SCAB_H modules and SCAB_H-enzyme fusions which were purified by size exclusion chromatography (SEC) because the engineered histidines into the protein caused them to remain anchored to the Ni-NTA resin. The purification protocol by Ni-NTA column was the

same as described above but with the difference that the protein was collected on the supernatant and the His-Tag remained attached to the Ni-NTA resin thus separating the two. The SCAB_H modules and SCAB_H-enzyme fusions were concentrated up to 2 mL and purified by SEC through a Superdex 75 (16/600) HL column using an ÄKTA prime plus fast protein liquid chromatography (FPLC) equipment. The chromatography was carried out in TRIS-HCl buffer (20 mM TRIS-HCl buffer pH 7.4) at a flow rate of 1 mL/min, and the chromatogram recorded at 280 nm.

Then, electrophoresis gels were used to confirm the molecular weight and the purity of the purified proteins, and the concentration was estimated based on the extinction coefficients calculated from their amino acid composition and the absorbance at 280 nm (Table 3.9).

Table 3.9. The Extinction coefficient of the proteins purified in this thesis. The purified proteins were: SCAB1_C, SCAB2_C, SCAB1_H, SCAB2_H, FDH@SCAB1_C, AlaDH@SCAB2_C, FDH@SCAB1_H and AlaDH@SCAB2_H, TRAP1-3, TRAP2-3-1, FDH1, ωTA2 and AlaDH3.

Protein	Extinction coefficient (M ⁻¹ .cm ⁻¹)
SCAB1 _C	38850
SCAB2 _C	33350
SCAB1 _H	47790
SCAB2 _H	47790
FDH@SCAB1 _C	91805
AlaDH@SCAB2 _C	56855
FDH@SCAB1 _H	96275
AlaDH@SCAB2 _H	69680
TRAP1-3	31665
TRAP2-3-1	48180
FDH1	51465
ωTA2	75330
AlaDH3	27390

SCAB_C modules were dialyzed into PBS buffer (150 mM NaCl, 50 mM phosphate buffer pH 7.4) and stored at -20 °C. SCAB_H modules were dialyzed into TRIS-HCl buffer (20 mM TRIS-HCl buffer pH 7.4) and stored at -20 °C. SCAB_C-enzyme fusions were dialyzed into PBS buffer (150 mM NaCl, 50 mM sodium phosphate buffer pH 7.4) and stored at 4 °C, to preserve their activity without requiring the presence of glycerol. SCAB_H-enzyme fusions were dialyzed into TRIS-HCl buffer (20 mM TRIS-HCl buffer pH 7.4) and stored at 4 °C, to preserve their activity without requiring the presence of glycerol. TRAP scaffolds were dialyzed into PBS buffer (150 mM NaCl, 50 mM sodium phosphate buffer pH 7.4) and stored at -80 °C. FDH1 and AlaDH3 tagged enzymes were dialyzed into PBS buffer (150 mM NaCl, 50 mM sodium phosphate buffer pH 7.4) and stored with 40 % glycerol at -80 °C, and ωTA2 tagged enzyme was dialyzed into TRIS-HCl buffer (150 mM NaCl, 50 mM TRIS-HCl pH 7.4) and stored with 10 % DMSO at -20 °C in order to preserve their activity (Figure 2.4).

3.3. MALDI-TOF mass spectrometry

MALDI-TOF measurements were performed using UltrafleXtreme III MALDI-TOF mass spectrometer with delayed extraction (Bruker) equipped with a pulsed nitrogen laser ($\lambda = 337$ nm). The samples were prepared in PB buffer (10 mM sodium phosphate buffer pH 7.4). 1 μ L of protein in a concentration range of 40-10 μ M was combined with 4 μ L of matrix. Then, this mixture was deposited on the MALDI-TOF plate and air dried. The preparation of the matrix solution was carried out by mixing 50:50 of acetonitrile-water and 0.1 % of TFA (Trifluoroacetic acid) with sinapinic acid matrix at 10 mg·mL⁻¹ final concentration¹⁰¹. All mass spectra were collected in positive reflection mode utilizing delayed extraction with a laser shot range of 50-100 and an excitation voltage of 20 kV (Figure 2.4).

3.4. Circular dichroism

Circular dichroism (CD)^{102,103} experiments were performed in PBS buffer (10 mM NaCl, 10 mM sodium phosphate buffer pH 7.4) at 10 μ M protein concentration using a Jasco J-815 spectrophotometer. CD spectra of the proteins was measured between 190-260 nm by averaging the spectra three times and were acquired in a 1 cm path length quartz cuvette. All CD spectra were recorded with a bandwidth of 1 nm in 1 nm increments and a mean time of 10 seconds.

Thermal denaturation curves were acquired by tracking the CD signal at 222 nm wavelength in the range of temperatures between 15 to 100 °C by recording data every 1 degree with digital integration time of 2 seconds. Thermal denaturation curves were performed in the same buffer and with the same protein concentration as in the CD experiments (Figure 3.4).

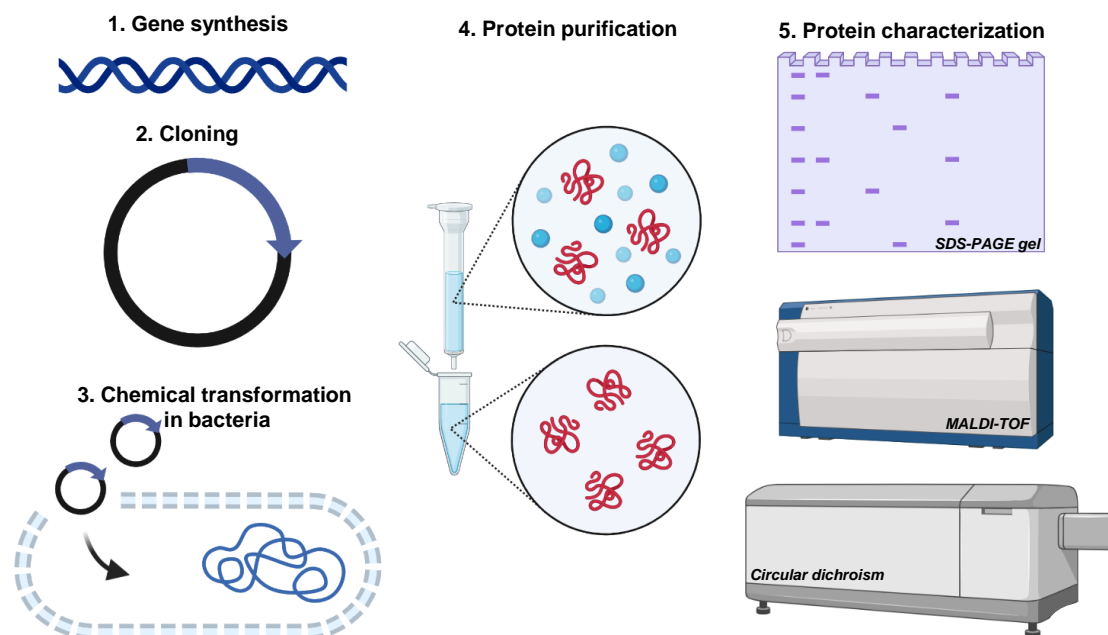


Figure 3.4. Production and primary characterization of recombinant proteins. 1. Gene of interest synthesis. 2. Recombinant plasmid cloning of the gene. 3. The plasmid is chemically transformed into bacteria. 4. Bacterial protein expression and purification. 5.

SDS-PAGE gels, MALDI-TOF, and circular dichroism were used to characterize the isolated proteins.

3.5. SCAB_C modules and SCAB_C-enzyme fusions assembly

The formation of orthogonal bio-bricks as intra-repeat interactions was stabilized with stapled di-sulfide bonds based on **SCAB_C modules**⁸². The process was performed in the solid state using a Ni-NTA resin and was based on the "head-to-tail-interactions" between the complementary scaffold modules SCAB1_C and SCAB2_C⁸⁷. The first step was activation of the cysteine at the C-terminal end of the SCAB1_C module. This was done by adding 10 µL of 1 M Dithiothreitol (DTT) (CAS number: 3483-12-3, Merck) into 1 mL volume of SCAB1_C module at 70 µM concentration in PBS buffer (150 mM NaCl, 50 mM sodium phosphate buffer pH 7.4) to a final activation concentration of 50 µM and incubating at room temperature for 20 min. Excess DTT was removed with the PD-10 column. Then, to avoid self-dimerization, incubation with 100 µL of excess aldrithiol (20 mM in DMSO) (CAS number: 2127-03-9, Merck) into 1 mL volume of SCAB1_C module at 50 µM concentration in PBS buffer (150 mM NaCl, 50 mM sodium phosphate buffer pH 7.4) to a final activation concentration of 20 µM was performed for 1 h at 40 °C⁹⁸. Excess aldrithiol was removed using a PD-10 column. Finally, 1 mL at 20 µM protein concentration in PBS buffer (150 mM NaCl, 50 mM sodium phosphate buffer pH 7.4) of neat SCAB1_C module was incubated with a Ni-NTA resin for 1 h at room temperature to promote solid-phase assembly. In turn, the cysteine at the N-terminal end of the SCAB2_C module was activated with DTT with the same methodology and buffers mentioned above. The only difference is the starting concentration of SCAB2_C module in order to incubate with SCAB1_C module in 1:10 of SCAB1_C:SCAB2_C molar ratio. Excess DTT was removed using the PD-10 column. In the next step, 1 mL at 200 µM protein concentration in PBS buffer (150 mM NaCl, 50 mM sodium phosphate buffer pH 7.4) of the cleaned SCAB2_C module was incubated with the previously immobilized SCAB1_C module to form on the "head-to-tail interaction" a di-sulfide covalent bond that clamps the interaction overnight at room temperature. After overnight incubation, excess non immobilized SCAB2_C module was cleaned up using PBS buffer (150 mM NaCl, 50 mM sodium phosphate buffer pH 7.4). Finally, the assembly formed by SCAB_C was eluted in 1 mL in PBS buffer with imidazole (150 mM NaCl, 300 mM imidazole, 50 mM sodium phosphate buffer pH 7.4).

However, in the case of the **SCAB_C-enzyme fusions** (FDH@SCAB1_C and AlaDH@SCAB2_C) the process was quite similar, but there were some differences. For example, incubation for di-sulfide bond formation was done with 0.5 mL at 50 µM protein concentration for each SCAB_C-enzyme fusions in PBS buffer (150 mM NaCl, 50 mM sodium phosphate buffer pH 7.4) and in solution at 4 °C overnight to promote assembly between these proteins of considerable dimensions. After overnight, 1 mL the assembly of the SCAB_C-fusions at 25 µM final concentration was incubated 1 h at room temperature in Ni-NTA resin in order to clean and purify the assembled proteins from the unassembled ones using PBS buffer (150 mM NaCl, 50 mM sodium phosphate buffer pH 7.4). Finally, the assembly formed with FDH/AlaDH@SCAB_C was eluted in 1 mL in PBS buffer with imidazole (150 mM NaCl, 300 mM imidazole, 50 mM sodium phosphate buffer pH 7.4). For the assembly FDH@SCAB1_C and AlaDH@SCAB2_C cysteines activation and clean-up processes were carried out using the same methodology described for SCAB1_C and

SCAB2_C modules. However, since the assembly process was done in solution, activation with aldrithiol was omitted (Figure 3.5, left).

3.6. SCAB_H modules and SCAB_H-enzyme fusions assembly

The formation of orthogonal bio-bricks as intra repeated interactions was stabilized with a coordination between **SCAB_H modules** with coordination histidines and metals^{84–86}. The process was performed in solution and was based on the metal driven assembly between complementary SCAB1_H and SCAB2_H scaffolding modules with designed coordination histidines and copper metal. The first step was the incubation of 0.25 mL of each SCAB_H module at 100 μM concentration in TRIS-HCl buffer (20 mM TRIS-HCl buffer pH7.4) and copper salt at 100 μM (CAS number: 7758-99-8, Merck) concentration for 1 h, at 50 °C and 850 rpm. Then, the 0.5 mL of the assembly through metal driven assembly at final concentration of 50 μM of each SCAB_H modules was clean and purified by SEC over a Superdex 75 (10/300) GL column using an ÄKTA prime plus FPLC equipment. The chromatography was carried out in TRIS-HCl buffer (20 mM TRIS-HCl buffer pH 7.4) at a flow rate of 0.5 mL/min, and the chromatogram recorded at 280 nm.

However, in the case of the **SCAB_H-enzyme fusions** (FDH@SCAB1_H and AlaDH@SCAB2_H) the synthesis process was quite similar, but there were some differences. For example, the incubation temperature of 0.25 mL SCAB-Histidine-enzyme fusions at 100 μM protein concentration in TRIS-HCl buffer (20 mM TRIS-HCl buffer pH 7.4). Due to the presence of enzymes this was decreased until 30 °C with the aim of avoiding the activity of the enzymes. Moreover, when the assembly was performed with a final volume of 0.5 mL at final concentration of 50 μM of each SCAB_H-enzyme fusion. Then, the assembly was clean and purified by SEC over a Superdex 200 (10/300) GL column using an ÄKTA prime plus FPLC equipment. The chromatography was carried out in TRIS-HCl buffer (20 mM TRIS-HCl buffer pH 7.4) at a flow rate of 0.5 mL/min, and the chromatogram recorded at 280 nm (Figure 3.5, middle).

3.7. Enzyme assembly onto TRAP proteins

The assembly between TRAP scaffold (**TRAP1-3**) and tagged enzymes (**FDH1 and AlaDH3**) is promoted by specific biomolecular recognition interaction⁹³. Before incubation, the His-tag of the tagged enzymes was removed by cleavage with TEV protease, whereas the His-tag of TRAP1-3 was not removed. The assembly was carried out in a systematic manner. Initially, AlaDH3 hexamers were incubated for 1 hour at 4 °C with TRAP1-3 monomers at a 1:6 stoichiometry to load all of the AlaDH3 monomers with one TRAP unit. When evaluating the monomeric units, FDH1 dimers were incubated 1 hour at 4 °C at a stoichiometry of one monomer per two TRAPs, resulting in an assembly with a stoichiometry of 6:3:1 (TRAP1-3:FDH1:AlaDH3), *i.e.* 1:1:1, when considering the monomeric units. 100 μL at 30 μM in PBS buffer 150 mM NaCl, 50 mM sodium phosphate buffer pH 7.4) of each assembly element were incubated yielding a total volume of 300 μL and a final concentration of 10 μM. Finally, the assemblies were purified by SEC through a Superdex 200 (10/300) GL column using an ÄKTA prime plus FPLC equipment. The chromatography was carried out in PBS buffer (150 mM NaCl, 50 mM sodium phosphate buffer pH 7.4) at a flow rate of 0.5 mL/min, and the chromatogram recorded at 280 nm (Figure 3.5, right).

The assembly between the TRAP scaffold (**TRAP2-3-1**) and the three tagged enzymes (**FDH1**, **ω TA2** and **AlaDH3**) was promoted by the specific biomolecular recognition interaction as the two-enzyme assembly. The incubation temperature was 4 °C and the assembly stoichiometry was 6:3:1.5:1 (TRAP1-3:FDH1: ω TA2:AlaDH3), *i.e.* 1:1:1:1, when considering monomeric units. 100 μ L at 40 μ M protein concentration of each assembly element in PBS buffer (150 mM NaCl, 50 mM sodium phosphate buffer pH 7.4) with a final assembly yield of 400 μ L and the concentration of the assembled elements was 10 μ M maintaining the 1:1:1:1 molar ratio for TRAP2-3-1:FDH1: ω TA2:AlaDH3. Finally, the assembly was purified by SEC through a Superdex 200 (10/300) GL column using an ÄKTA prime plus FPLC equipment. The chromatography was carried out in PBS buffer (150 mM NaCl, 50 mM sodium phosphate buffer pH 7.4) at a flow rate of 0.5 mL/min, and the chromatogram recorded at 280 nm.

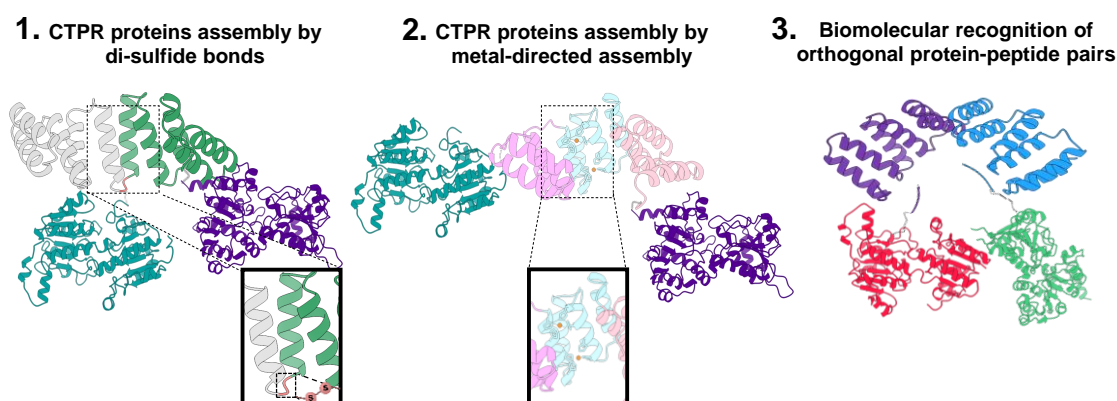


Figure 3.5. The three assembly techniques are shown schematically. The first and second ways rely on CTPR proteins fused directly to enzymes. Prior design results in a covalent link created by di-sulfide bonding stapled by cysteines and assembly driven by histidines as coordinating residues with the metal copper. The third technique relies on the biomolecular recognition of TRAP scaffolds and peptide-containing enzymes to assemble. The affinity of peptides for TRAP scaffolds promotes this assembly.

3.8. Determination of the apparent Hydrodynamic radius (R_h)

The Fluidity One W system is based on the diffusion of labeled molecules. Fluidity-One W has as its main objective to calculate metrics related to the size of the fluorescently labeled molecules. Using a chip as an intermediary, 5 to 10 μ L of your sample is deposited on the chip and within 10-15 minutes you obtain the apparent hydrodynamic radius of the analyzed systems by tracking the fluorescence of the labeled molecule in addition to diffusion and employs a unique diffusion-based technique (Figure 3.6).



Figure 3.6. Fluidity One W system. The Fluidity One W system (www.fluidic.com/product/fluidity-one-m/) is based on microfluidic diffusional sizing technology. Samples are deposited on a microfluidic chip and then introduced into the slot of the Fluidity One W system. After about 10-15 minutes, parameters such as apparent Rh are obtained.

3.8.1. SCAB_H assembly

In the SCAB_H assembly, an analysis by microfluidic diffusional sizing technology by Fluidity One W system⁹⁹ was performed to study the increase in apparent Rh by playing with the molar ratios between two SCAB_H modules. 50 μL of SCAB_H assemblies (25 μL of each SCAB_H module) were prepared with different SCAB_{1H} protein concentrations 20, 30, 50, 50, 100 and 150 μM and fixed SCAB_{2H} concentration 50 μM , which means 0.4:1, 0.6:1, 1:1, 1:2 and 1:3 molar ratio of SCAB_{2H}:SCAB_{1H}, respectively. The assemblies were done in TRIS-HCl buffer (20 mM TRIS-HCl buffer pH 7.4). For preparing assemblies and measure by microfluidic diffusional sizing SCAB_{2H} at 70 μM protein concentration in TRIS-HCl buffer (20 mM TRIS-HCl buffer pH 7.4) was incubated for 1 h with a solution of Alexa-Fluor 647 (AF647) dye in DMSO at a protein:dye molar ratio of 1:3, in the dark with shaking (25 rpm) at 25 °C. Excess dye was removed by gel filtration through a PD-10 column. Then, after labeled SCAB_{2H}, SCAB_{2H}:SCAB_{1H} modules in different molar ratios were incubated for 1 h at 50 °C with 100 μM CuSO₄ salt in TRIS-HCl buffer (20 mM TRIS-HCl buffer pH 7.4). Excess metallic Cu was washed off with PD-10 column. Finally, 10 μL of the assemblies were used to determine the apparent Rh of the different assembled systems, fluorescence tracking of the SCAB_{2H} module was measured at a concentration of 50 μM . All apparent Rh measurements were taken in duplicate for each assembled system.

3.8.2. AlaDH/FDH@TRAP1-3 assembly

The apparent Rh of the TRAP scaffold, free enzymes, and assembled enzymes systems was measured by microfluidic diffusional sizing technology using the Fluidity One W

system¹⁰⁴. 1 mL of TRAP1-3 scaffold at 30 μM protein concentration in PBS buffer (150 mM NaCl, 50 mM sodium phosphate pH 7.4) was labeled with AF647 as fluorescent reporter to determine the diffusion and thus the apparent Rh. 1 mL of TRAP1-3, FDH1 and AlaDH3 at 10 μM protein concentration in PBS buffer (150 mM NaCl, 50 mM sodium phosphate pH 7.4) were incubated for 1 h with A647 dye solution in DMSO at 1:3 protein:dye molar ratio, in darkness with agitation (25 rpm) at 25 °C. The excess of the dye was removed by gel filtration through a PD-10 column. First, the apparent Rh of the free tagged enzymes and 10 μL of TRAP1-3 scaffold was analyzed at 2 μM protein concentration. Then, in order to determine the apparent Rh of the assemblies, unlabeled tagged enzymes were incubated sequentially with labeled TRAP1-3 scaffold. First AlaDH3 was incubated for 1 h at 4 °C and then FDH1 was incubated for 1 h at 4 °C, *i.e.*, 2 h in total in a 1:1:1 molar ratio for TRAP1-3:FDH1:AlaDH3 (protein concentration 10 μM) in PBS buffer (150 mM NaCl, 50 mM sodium phosphate pH 7.4). Then, the assemblies were 1:5 diluted and tracking 10 μL the apparent Rh of the different scaffolded enzyme systems was measured by tracking the fluorescence of the TRAP1-3 scaffold at 2 μM concentration. All apparent Rh measurements were taken in duplicate for each element, free or assembled.

3.9. Size determination by size exclusion chromatography (SEC)

In order to analyze the differences in the elution peaks volume of the assembled and non-assembled elements, the SCAB modules and the SCAB-enzyme fusions with cysteines and histidines were analyzed by SEC¹⁰⁰. Once the individual elements were analyzed, the increments were corroborated in relation to the elution peaks volume of the four assemblies obtained: SCAB_C, SCAB_H, FDH/AlaDH@SCAB_C and FDH/AlaDH@SCAB_H. Assembled and non-assembled SCAB modules were injected into SEC on a Superdex 75 (10/300) GL size exclusion column, while assembled and non-assembled SCAB-enzyme fusions were injected into SEC on a Superdex 200 (10/300) GL size exclusion column. Runs were performed at 0.5 mL/min in PBS buffer (150 mM NaCl, 50 mM PB buffer pH 7.4) and TRIS-HCl buffer (20 mM TRIS-HCl buffer pH 7.4) for the cysteine and histidine protein variants, respectively, at 4 °C. The volume used for all samples analyzed by SEC was 500 μL . The protein concentration of each SCAB_C modules and SCAB_C-enzyme fusions was 25 μM , while the protein concentration of each SCAB_H modules and SCAB_H-enzyme fusions was 50 μM .

3.10. Inductively coupled plasma mass spectrometry (ICP-MS)

100 μL of each of the SEC-purified SCAB_H and FDH/AlaDH@SCAB_H assemblies sets at a concentration of 13.07 and 11.65 μM respectively, were mixed with 300 μL of 37 % HCl and the resulting suspension was sonicated in a water bath for 45 min at 60 °C. Finally, 2700 μL of distilled water at pH 7.4 was added obtaining a final volume of 3000 μL ¹⁰⁵. The Cu metal concentration of the assembled systems previously purified by SEC was determined by measuring the sample by ICP-MS iCAP-Q from Thermo ScientificTM. ICP-MS is a combination of mass spectrometry using an inductively coupled plasma which ionizes samples. Once this process has been carried out, the sample creates atomic ions and small polyatomic ions, which are then detected.

3.11. Enzymatic activity measurements

Enzyme activity, enzyme kinetic parameters, L-Alanine biotransformation and competitive side reaction catalytic assay have been determined using the same protocols in all three scaffold strategies: head-to-tail interactions stapled by a di-sulfide bonds and metal-driven interactions, and biomolecular recognition. Therefore, for easy identification the enzymes in these assays are referred to as FDH, ω TA and AlaDH.

The activity of free and scaffolded enzymes was measured in solution¹⁰⁶. To determine the AlaDH activity, the consumption of NADH was measured by the decrease in absorbance at 340 nm under the following reaction conditions: 0.5 mM NADH, 75 mM pyruvate, 500 mM ammonium chloride in 25 mM potassium phosphate pH 8.0 at 0.2 μ M AlaDH. To determine the FDH activity, the production of NADH was measured by the increase in absorbance at 340 nm under the following reaction conditions: 1 mM NAD⁺, 100 mM sodium formate, in 25 mM sodium phosphate buffer pH 7.0 at 2 μ M FDH. All the measurements were performed in 200 μ L reaction volume during 30 min, at a 340 nm wavelength and at 30 °C using Synergy H1 Hybrid Multi-Mode Microplate Reader from BioTeK Instrument in 96-well UV-Vis transparent plates. To determine the activity of the ω TA, the conversion of methylbenzylamine (FEA) to acetophenone was measured in 200 μ L by monitoring an increase in the absorbance during 15 minutes at 245 nm under the following reaction conditions: 2 mM pyruvate, 0.1 mM PLP (Pyridoxal 5'-phosphate monohydrate) and 2 mM FEA in acetonitrile in 200 mM HEPES buffer pH 8.0 at 0.5 μ M ω TA. The slope of absorbance as a function of time was calculated by a linear fit of the first time points of each reaction to calculate the enzyme activity units according to the Lambert-Beer equation. For FDH and AlaDH, the NADH concentration was calculated using the extinction coefficient (ϵ) $\epsilon_{\text{NADH-340 nm}} = 6220 \text{ M}^{-1} \times \text{cm}^{-1}$. One unit of FDH is defined as the amount of enzyme needed to produce 1 μ mol of NADH per minute under the above given conditions. One unit of AlaDH is defined as the amount of enzyme needed to consume 1 μ mol of NADH per minute under the above given conditions. For the ω TA, we used the extinction coefficient (ϵ) of acetophenone; $\epsilon_{\text{-245 nm}} = 12000 \text{ M}^{-1} \times \text{cm}^{-1}$. One unit of ω TA is defined as the amount of enzyme needed to produce 1 μ mol of acetophenone per minute under the above given conditions. All enzymatic activity measurements were made in triplicate and the standard deviation was calculated from these replicates. The specific activity was calculated normalizing the activity units per the enzyme concentration in the assay ($\text{U} \cdot \text{mg}^{-1}$)

$$\text{Specific enzymatic activity } (\text{U} \cdot \text{mg}^{-1}): \left(\frac{\text{Enzyme activity } (\text{U} \cdot \text{mL}^{-1})}{\text{Protein concentration } (\text{mg} \cdot \text{mL}^{-1})} \right)$$

3.12. Enzyme kinetic parameters

Kinetic parameters: Michaelis-Menten constant (K_M), maximum rate (V_{max}), turnover number, (k_{cat}) and catalytic efficiency (k_{cat}/K_M) of free and scaffolded FDH and AlaDH were determined spectrophotometrically by Synergy H1 Hybrid Multi-Mode Microplate Reader from BioTeK Instrument in 96-well UV-Vis transparent plates according to the enzymatic activity assays described above. Different concentration ranges of the cofactors NADH (from 0 to 0.66 mM) and NAD⁺ (from 0 to 8.33 mM), pyruvate (from 0.9 to 200 mM), formate (from 0 to 62.5 mM), and ammonium chloride (500 mM) were evaluated to calculate the

K_M and V_{max} . Two Michaelis-Menten curves were acquired for the FDH enzyme. In the first one the concentration of formate varied and the concentration of NAD^+ was fixed (1 mM). However, in the second one it was NAD^+ that varied in concentration and formate maintained its concentration (100 mM). The buffer used was 25 mM sodium phosphate buffer pH 7.0. Two Michaelis-Menten curves were also obtained for the AlaDH enzyme. In the first one the pyruvate concentration varied and the NADH concentration was fixed (0.5 mM). In the second one it was NADH that varied in concentration and pyruvate maintained its concentration (75 mM). The buffer used was 25 mM potassium phosphate buffer pH 8.0. Experimental data were fitted by the Michaelis-Menten equation and are reported in specific enzymatic activity values ¹⁰⁷:

$$V = \left(\frac{V_{max} \times [S]}{K_M + [S]} \right)$$

where V is the initial velocity, $[S]$ is the substrate concentration, V_{max} the maximum reaction velocity, and K_M the Michaelis-Menten constant. The activity of AlaDH, in which pyruvate is the substrate, since the data does not fit well to a simple Michaelis-Menten equation it was fitted using the following substrate inhibition equation ¹⁰⁸:

$$V = V_{max} \frac{[S]}{[S] + K_M \left(1 + \frac{I}{K_i} \right)}$$

where V is the initial velocity, V_{max} is the maximum velocity, K_M is the substrate binding constant (Michaelis-Menten substrate affinity constant), and K_i the inhibitor binding constant.

The standard deviation reported for the kinetic parameters, K_M and V_{max} , were calculated from the average of three replicates. The turnover numbers (k_{cat}) were obtained using the following equation:

$$k_{cat} = V_{max}/[\text{Enzyme}]$$

The ratio k_{cat}/K_M defines the catalytic efficiency of the systems.

3.13. Biotransformation of L-Alanine

L-Alanine was synthesized in batch incubating free and scaffolded biocatalysts composed by AlaDH and FDH at two different AlaDH:FDH molar ratios 1:1 and 1:8, which correspond to of 0.28:0.28 μM (monomer based), and 0.28:2.25 μM (monomer based) AlaDH:FDH protein concentration in 1 mL of the reaction mixture composed by 75 mM pyruvate, 100 mM sodium formate, 500 mM ammonium chloride and 0.5 mM NADH in nano pure water, pH 7.4. In the scaffold strategies based on head-to-tail and metal-driven interactions only the 1:1 molar ratio for AlaDH:FDH was analyzed. Reactions were incubated under orbital agitation at 500 rpm and 25 °C for 1, 2, 4, 8, and 24 h. The reaction was stopped, and the L-Alanine was collected at each time point by passing the samples through an ultrafiltration unit Amicon Ultra-0.5 Centrifugal Filter Units, which were centrifuged 30 min at 14000 rpm. The conversion degree of L-Alanine was confirmed by chiral derivatization with Marfey's reagent ¹⁰⁹ and analyzed by High-performance liquid chromatography (HPLC). Briefly, 20 μL of 1:10 diluted reaction samples were mixed with 8 μL of 1 M sodium bicarbonate and

20 μL of 15 mM Marfey's reagent (Cat. 48895, Thermo Scientific) in acetone and incubated for 1 h at 50 °C and 400 rpm. Then, the derivatized reaction was stopped by the addition of 8 μL of 2 M HCl and then centrifuged at 10000 rpm for 15 min. In addition, the supernatant was filtered to perform the HPLC analysis. Derivatized samples were analyzed in a HPLC Agilent Technologies 1120 Compact LC, with an EC-C18 2.7 μm column (4.6 x 100 mm, Agilent) with the mobile phases A (0.1 % TFA in water) and B (Acetonitrile) at 1 $\text{mL}\cdot\text{min}^{-1}$ flow rate. Analytes were detected at 340 nm and eluted with the following gradient: starting from 90-80 % A from 0-17 min, then from 80-60 % A from 17-20 min, the mobile phase was maintained at 60 % from 20-30 min, restored to initial conditions 90 % A in 1 minute and kept the mobile phase at 90 % from 31-40 minutes. The conversion degree of L-Alanine was calculated by fitting the peak's area with a calibration curve.

3.14. Fluorescence anisotropy-based binding assay

The procedures described below only correspond to the assembly strategy based on biomolecular recognition. Therefore, the enzymes are named FDH1, ωTA2 and AlaDH3 in order to refer to the presence of the peptide.

The binding of NADH cofactor to TRAP1-3 scaffold was conducted in PBS buffer (150 mM NaCl, 50 mM PB pH 7.4) in a 10 x 10 mm path-length cuvette at 25 °C at 1 M NADH concentration. After 5 minutes of equilibration, increasing quantities of the TRAP1-3 scaffold (from 0 to 100 μM TRAP1-3 concentration) were added to the NADH solution (100 μL at 1 μM NADH concentration), and the fluorescence anisotropy was measured. Experiments on fluorescence anisotropy were carried out using a Fluorometer NIR fluorescence spectrophotometer outfitted with excitation and emission polarizers. Fluorescence anisotropy was measured with excitation at 340 nm and emission at 463 nm using slits of 6 nm.

Anisotropy was calculated using the equation G-factor: adjustments were performed to account for the difference in transmission efficiency of the two emission channels.

$$G - factor = \frac{I_{HV} - I_{B,HV}}{I_{HH} - I_{B,HH}}$$

where I_{HV} is the vertical emission (0°) of a standard solution with excitation in horizontal orientation (90°), I_{HH} is the horizontal emission of a standard solution with excitation in horizontal orientation, $I_{B,HV}$ is the vertical emission of a blank solution with excitation in horizontal orientation, and $I_{B,HH}$ is the horizontal emission of a blank solution with excitation in horizontal orientation. PBS buffer (150 mM NaCl, 50 mM PB pH 7.4) was used as a blank solution and 1 μM NADH as a standard solution.

The equation for anisotropy (r) includes the G-factor for excitation at vertical orientation (0°) is:

$$r = \frac{(I_{VV} - I_{B,VV}) - G(I_{VH} - I_{B,VH})}{(I_{VV} - I_{B,VV}) + 2G(I_{VH} - I_{B,VH})}$$

where G is the G-factor, I_{VV} and I_{VH} are the sample's vertical and horizontal emission, respectively, and $I_{B,VV}$ and $I_{B,VH}$ are the intensity of emission of the blank with the emission polarizer in vertical and horizontal orientations, respectively.

To calculate the percentage of NADH bound at different concentrations it was used the following equation ¹¹⁰:

$$\text{Binding \%} = \frac{r - r_f}{r_b - r_f}$$

where r represents the measured anisotropy for NADH at any TRAP1-3 scaffold concentration, r_f represents the anisotropy of free NADH, and r_b represents the anisotropy of NADH bound to the TRAP1-3 scaffold in the plateau area of the binding curve. GraphPad Prism 9 software was used to fit the data to a One Site-Specific binding model.

3.15. Competitive side reaction catalytic assay

A competition test ¹¹¹ was designed to assess a potential substrate channeling effect ⁵⁴. Pure NADH oxidase (NOX) from *Thermus thermophilus* ¹⁰⁹ was added as a competitor to the multi-enzymatic reaction catalyzed by the scaffolded FDH1:AlaDH3 enzymes. NOX as AlaDH3 required NADH as a cofactor, therefore the activity of NOX will be a reporter of the capability NADH accumulation as NOX will use this cofactor as substrate concomitantly and stoichiometrically producing H_2O_2 . For the competitive catalytic assay, free and scaffolded FDH1 at 0.18 μM , AlaDH3 at 0.18 μM , and an excess of NOX (4.4 μM) were mixed in 200 μL of reaction mixture composed of 75 mM pyruvate, 100 mM sodium formate, 500 mM ammonium chloride, 0.5 mM NADH, 0.15 mM FAD^+ , 0.1 mg/mL HRP, and 0.05 mM Amplex Red (AR) in pure nano water, pH 7.4. The oxidation of the cofactor NADH to NAD^+ related to the activity of both AlaDH3 and NOX was measured at 340 nm. The activity of NOX was determined by the H_2O_2 formation through a coupled assay using Horseradish peroxidase (HRP) and the oxidation of Amplex red as reporter by monitoring the absorbance of resorufin at 560 nm. The specific enzymatic activity of the reactions at each wavelength were determined spectrophotometrically using Synergy H1 Hybrid Multi-Mode Microplate Reader from BioTek Instrument in 96-well UV-Vis transparent plates. The extinction coefficient to determine the specific enzymatic activity related to NADH oxidation is (ϵ) $\epsilon_{\text{NADH-340 nm}} = 6220 \text{ M}^{-1} \times \text{cm}^{-1}$, whereas the extinction coefficient to determine the specific enzymatic activity related to H_2O_2 formation which is related to amplex red oxidation (ϵ) $\epsilon_{\text{Amplex red-560 nm}} = 54000 \text{ M}^{-1} \times \text{cm}^{-1}$.

3.16. Isotope enrichment and dilution assay: deuterated and non-deuterated L-Alanine product formation

Deuterated and non-deuterated L-Alanine were synthesized in batch incubating free and scaffolded AlaDH3 and FDH1 at 1:1 molar ratio (monomer-based) at 0.18 μM protein concentration of each enzyme in 1 mL of the reaction mixture composed by 75 mM pyruvate, 75 mM deuterated sodium formate, 25 mM sodium formate, 500 mM ammonium chloride, 0.5 mM NADH, 4.4 μM NOX, and 0.15 mM FAD^+ in nano pure water, pH 7.4. Reactions were maintained under orbital agitation at 500 rpm and 25 °C for 1, 2, 4, 8, and 24 h. Reaction were stopped by tangential filtration at 14000 rpm using an Amicon Ultra-

0.5 centrifugal filter units for 30 minutes. The reaction samples at different points were diluted 10 times and then derivatized by Dansyl method¹¹² prior analysis by UPLC-MS (Ultra-Performance Liquid Chromatography-Mass spectrometer). UPLC was performed in a Waters ACQUITY UPLC system with a Acquity BEH C18 column (100 x 2,1 mm / 1,7 μm). The gradient elution solvents were A (100 mM Ammonium formate in Water) and B (Acetonitrile) at a flow rate of 300 $\mu\text{L}\cdot\text{min}^{-1}$, with the following gradient 80 % A, up to 1 % A for 28 minutes and back to 80 % A for the remaining two minutes for a total duration of 30 minutes. The conversion degree was determined for deuterated or non-deuterated L-Alanine calculating the abundance of both hydrogen isotopes in the analyzed systems by mass spectrometry detection carried out using a time-of-flight mass spectrometer (ESI-TOF) LCT Premier XE from Waters (Milford, MA, USA) with an electrospray ionization source, working in positive/W mode. The MS range acquired was between m/z 50–1000. The capillary and cone voltages were set at 3000 and 50 V, respectively. Desolvation gas temperature was 300 °C and source temperature was 120 °C. The desolvation gas flow was set at 600 $\text{L}\cdot\text{h}^{-1}$ and cone gas flow was set at 50 $\text{L}\cdot\text{h}^{-1}$. For quantification and data analysis, Masslynx v4.1 software was used (Waters, Milford, MA, USA). Dansyl derivatized L-Alanine detected with a mass of 323 $\text{g}\cdot\text{mol}^{-1}$ and derivatized deuterated L-Alanine detected with a mass of 324 $\text{g}\cdot\text{mol}^{-1}$.

3.17. Fluorescence Confocal Microscopy

For imaging the distribution of the enzymes and the potential co-localization on solid beads fluorescence confocal microscopy was used. The FDH1 and AlaDH3 tagged enzymes at 10 μM protein concentration and in PBS buffer (NaCl 150 mM, sodium phosphate buffer 50 mM pH 7.4) were labeled by mixing protein solutions in PBS buffer (NaCl 150 mM, sodium phosphate buffer 50 mM pH 7.4) with AF647 and Alexa-Fluor 488 (AF488) dye solutions in DMSO (1:3 molar ratio of protein:dye), respectively. The AF647 and AF488 protein labeling kits belong to Jenna Bioscience with NHS-Ester reactivity. The reactions were incubated 1 h in darkness with agitation at 25 °C. The excess of dyes was removed by gel filtration through a PD-10 column. Then, the labeled tagged enzymes were assembled onto the TRAP1-3 scaffold previously immobilized on two carriers functionalized with cobalt chelates (from Abts beads supplier) and cobalt chelates, positively charged amine groups and aldehydes at its surface, respectively. Finally, the complexes were immobilized on both carriers for 1 h in gentle agitation (25 rpm) at 25 °C. A suspension of the beads in sodium phosphate buffer 50 mM pH 7.4 (1:200) was analyzed by confocal microscopy using a Zeiss LSM 510 microscope by recording the red (λ_{ex} : 488 nm and emission filter LP505 for AF488) and green (λ_{ex} : 633 nm and emission filter LP650 for AF647) channels¹¹³. Confocal image processing was performed using Image J (FIJI).

3.18. L-Alanine synthesis with heterogeneous biocatalysts

Immobilization of free and scaffolded enzyme systems on a solid support was carried out using tri-functional carrier³³. First, 95 mg of tri-functional carrier was equilibrated with 950 μL of PBS buffer (150 mM NaCl, 50 mM potassium phosphate buffer pH 8.0) for 5 min at 25 °C in a 1 mL unpacked column. Then, 950 μL of the free or scaffolded FDH1:AlaDH3 systems at 1:1 molar ratio (monomer-based) with of each enzyme concentration of 2.8 μM , were incubated for 4 h at 4 °C with gentle agitation (40 rpm). After incubation, five

washes of the resin with PBS buffer (150 mM NaCl, 50 mM potassium phosphate buffer pH 8.0) were performed. Next, 950 μ L of 1 M glycine was added to remove any free aldehyde in the resin and incubated overnight at 4 °C with gentle agitation (40 rpm). After overnight incubation the resin was washed 5 times with PBS buffer (150 mM NaCl, 50 mM potassium phosphate buffer pH 8.0). Subsequently, the immobilization yield of both systems was determined by in a final volume of 200 μ L the specific enzymatic activity from initial, non-immobilized and immobilized FDH1 and AlaDH3 at 2 μ M and 0.2 μ M protein concentration, respectively. This measurement was performed spectrophotometrically by Synergy H1 Hybrid Multi-Mode Microplate Reader from BioTeK Instrument in 96-well UV-Vis transparent plates (as previously described in Enzymatic activity measurements section). Finally, reusability of the immobilized systems was assessed. To that aim, 50 mg of the immobilized biocatalyst (with FDH1:AlaDH3 free and scaffolded at 1:1 ratio) were incubated with a 450 μ L mixture composed by 500 mM ammonium chloride, 75 mM pyruvate, 100 mM sodium formate and 0.5 mM NADH in nanopure water at pH 7.4 to achieve a final enzyme concentration of each enzyme of 0.28 μ M for 24 h at 25 °C at 500 rpm. The reaction was stopped by tangential filtration at 14000 rpm using an Amicon Ultra-0.5 centrifugal filter units for 30 minutes. This was followed by 1 wash with 450 μ L of 25 mM potassium phosphate buffer pH 8.0. The same process was repeated for 5 consecutive cycles (24 hours each cycle) to analyze the reusability of the systems. The conversion degree of L-Alanine upon each cycle was determined by HPLC and calculated by fitting the peak's area with a calibration curve.

3.19. Benzylamine synthesis catalyzed by scaffolded three-enzyme multi-enzymatic system

The production of benzylamine by the scaffolded three-enzyme system was analyzed by HPLC. Benzylamine was synthesized in batch incubating of each enzyme concentration of 10 μ M the free and scaffolded biocatalysts (FDH1, ω TA2 and AlaDH3 at 1:1:1 molar ratio monomer-based) in 400 μ L of reaction mixture composed by 50 mM pyruvate, 500 mM ammonium formate, 0.1 mM PLP, 10 mM benzaldehyde and 0.5 mM NADH in nano pure water. Reactions were incubated under orbital agitation at 500 rpm and 25 °C for 2, 8, 24, and 48 h. Reaction were stopped, and reaction crude were collected at each time point by centrifuging 30 min at 14000 rpm in Amicon Ultra-0.5 Centrifugal Filter Units. Samples were directly analyzed using Agilent Technologies 1120 Compact LC HPLC with a 5 μ m Ultrabase C18 column (4.6 x 250 mm, PurpleSeries) with mobile phases A (0.1 % TFA in water) and B (acetonitrile) at a flow rate of 0.9 mL \cdot min⁻¹. The analytes were detected at 245 nm and eluted with the following gradient: 80-65 % A from 0-10 min, then 65 % A from 10-20 minutes, and finally then restoring the initial conditions in 10 min followed by a 10 min equilibration at these same initial conditions. The degree of conversion of benzylamine was calculated from the peak area by using a calibration curve.

**Chapter 4.
Engineering Bio-brick protein
scaffolds for enzyme assembly**

4.1. Introduction

4.1.1. Enzyme assembly via biomacromolecules: fusion proteins

There is a growing need to develop new scaffolds for the assembly of enzymes in order to organize biomacromolecules in space and facilitate cascade reactions. The demand for efficient and adaptive synthetic approaches¹¹⁴ has made enzyme assemblies increasingly important⁴². Protein scaffolds have demonstrated their effectiveness in promoting enzyme assembly, enhancing the catalytic efficiency of multi-enzyme systems, and addressing diffusion challenges³⁴ that are commonly encountered with traditional supports like solid carriers³⁴. Fusion proteins are one of the preferred alternatives for enzyme assembly as they can be synthetically altered to promote improved characteristics or novel functionalities⁴. Enzymes are often bound together to adequately provide or recycle cofactors such as NADH and NAD(P)H, making fusion enzyme engineering a promising technique for enzyme cascade, especially when *in-situ* cofactor recycling is needed^{115,116} (Figure 4.1).

In recent years, more complex systems have been developed to enhance enzyme performance. Among the most prominent are the engineering of enzyme active sites or surfaces and assembly by fusion proteins¹¹⁶. Although the catalytic efficiency of the process can be significantly increased by bringing two enzymes together, fusion proteins often face drawbacks like aggregation into inactive enzyme clusters or reduction of enzyme intrinsic kinetics¹⁸. These limitations can be partially overcome by more sophisticated methods based on post-translational fusion proteins enabled by synthetic protein scaffolds^{45,117}. The most significant advantages of having the enzymes in protein-based scaffolds is enzymes are fused upon translation minimizing the chances of aggregation underlying large artificial polypeptides^{42,43}. Furthermore, the multi-enzyme systems based on protein-based scaffolds can be used both *in-vivo* and *in-vitro*. and therefore, a control of nanometric spatial organization of the enzymes is achieved⁴⁵ (Figure 4.1).

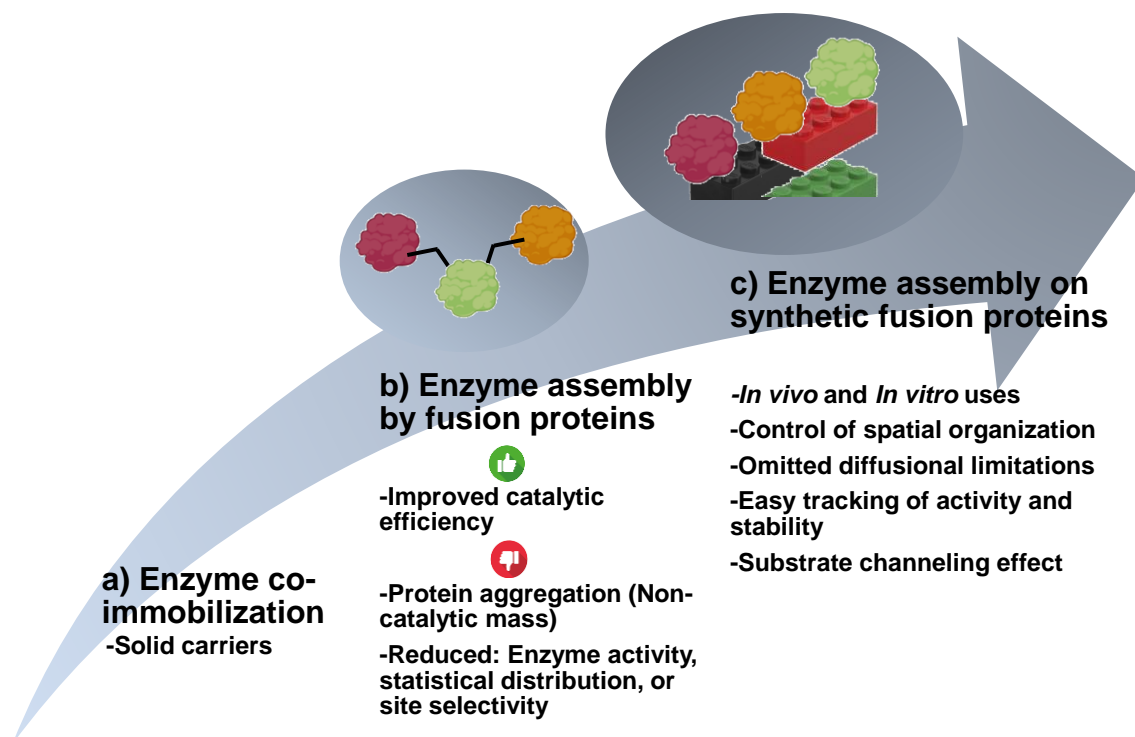


Figure 4.1. From enzyme co-immobilization to assembled enzymes on synthetic fusion proteins. a) Enzyme co-immobilization on solid carriers to obtain enzymatic activity and stability. b) Increased catalytic efficiency in assemble multi-enzyme systems formed by biomacromolecules (fusion proteins) is the most outstanding advantage. Some of the disadvantages of assemble multi-enzyme systems formed by fusion proteins: protein aggregation and reduction of enzymatic activity, statistical distribution, or site selectivity. c) Enzyme assembly on synthetic fusion proteins: *In vivo* and *In vitro*, controlled spatial organization, diffusional limitations omitted, easy control of enzyme activity and stability, and substrate channeling effect.

Several synthetic protein scaffolds have been developed to assemble enzymes to enhance the one-pot transformations of alcohols and aldehydes into amines ⁴⁴. Zeballos et al. ³² assembled two enzymes on a synthetic protein scaffold which, in turn, is immobilized on a solid carrier, a porous agarose microbead. Regarding the synthetic protein scaffold, the enzymes are assembled by a cellulosome-based system, the dockerin-cohesin pair. The association between cohesin and dockerin serves as the foundation for the inclusion of individual enzyme subunits into the cellulosome complex ¹¹⁸. Through a more sophisticated approach, Ledesma *et al.* have developed a protein scaffold based on Tetratricopeptide repeat affinity proteins (TRAPs) that spatially arrange up to three enzymes, locally sequestering the required cofactors, resulting a more productive cascade than its free counterpart ¹¹⁹. The aforementioned protein complexes are based on reversible protein-protein interactions where covalent bonds are not involved. In contrast, Zhang et al. ⁴⁴ exploited another widely exploited used a protein domain-based protein synthetic scaffold, the SpyTag-SpyCatcher system, to covalent and irreversibility assemble two orthogonal dehydrogenases The SpyTag-SpyCatcher protein scaffold ⁶⁹ promotes a dual enzyme cascade for chiral amine synthesis and demonstrates

that scaffolding the cascade decreases the time necessary to accomplish final conversions when compared to a free enzyme system⁶⁹.

4.1.2. CTPR proteins as scaffolding units

As an alternative to the described protein-protein interactions, the consensus tetrapeptide repeat (CTPR) module^{74,82} was selected for the design and fabrication of a novel set of modular synthetic protein scaffolds (SCABs) that can assemble multi-enzyme systems. CTPR units present numerous characteristics that make them ideal scaffolds for this application: 1) modular nature¹²⁰; 2) notable stability¹²¹; 3) tolerance to mutations; 4) structure defined by few conserved residues within their sequence⁷⁴; 5) intrinsic self-assembly properties to form linear protein arrays^{87,122}; and 6) extended structure that displays a large surface area to volume ratio¹²³. These features have allowed previous demonstrations of the potential of engineered CTPR proteins as scaffolds template different small molecules and nanoelements, such as, photoactive molecules⁸³, single-walled carbon nanotubes⁸⁴, electroactive clusters⁸⁵, and gold nanoparticles⁸⁶. Moreover, the CTPR proteins are more resistant to the destabilizing effects of mutations because they are thermodynamically more stable than their natural counterparts, TPRs^{81,82}. They have also been found to be more thermodynamically stable than another synthetic protein based on TPR domains, TRAPs (Tetratricopeptide repeat affinity proteins). The melting temperature (T_M) at which the proteins begin to lose their secondary structure is about 60 °C for CTPRs comprise of 3 CTPR repeats⁸⁰ and is to about 50 °C for TRAPs^{93,124}. Equally noteworthy, TRAP proteins⁹³ have an architecture that is nearly identical to that of their natural counterpart, TPR. As a result of their more constrained design, they are less versatile in terms of application than CTPR proteins^{123,125}. Furthermore, it is possible to assemble identical CTPR units into linear arrays of up to hundreds of nanometers because to the well-researched CTPR inter-repeat packing interface. Additionally, the insertion of certain cysteines at the N- and C- terminal ends of the units allows the development of disulfide bonds that stitch those connections, creating covalently connected linear nanofibers that are incapable of dissolving spontaneously in water solution⁸⁷. Another feature of CTPR proteins is that to encode larger CTPR proteins with super helical structures, CTPRs can be tandemly repeated or coupled to other peptides or proteins such as enzymes. As a result, CTPR proteins are a highly suitable choice-candidate for utilizing as protein scaffold to enzyme assemblies (Figure 4.2).

The CTPR unit is a helix-turn-helix motif made up of 34 amino acids, only 8 of which are necessary to enable proper structural folding. Because there are few conserved residues, it is easier and more reproducible to adapt functional mutations to provide the protein with the desired features. For instance, the addition of metal-binding residues, like histidines and cysteines, allow the coordination of metal ions and the subsequent stabilization of novel nanomaterials within the CTPR proteins^{74,123} (Figure 4.2).

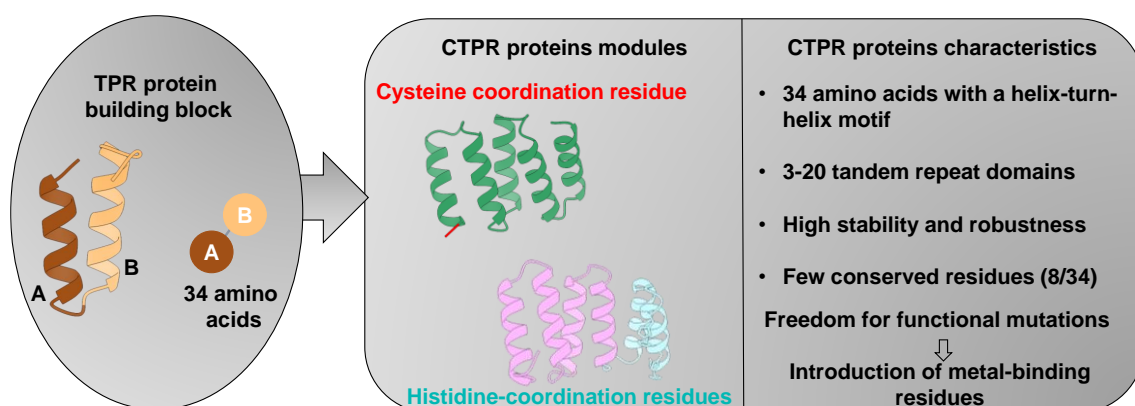


Figure 4.2. Schematic representation of CTPR proteins and their natural counterpart, TPR protein domain. On the left is shown a module of the natural TPR domain consisting of 34 amino acids that encode a helix-turn-helix motif. On the right, the two CTPR modules with their respective coordination sites; cysteine-based coordination site (in red) and histidine-based coordination site (in cyan). The most important features of the CTPR proteins are listed, highlighting the most important one: versatility; that is, freedom for functional mutations, for example through the introduction of metal-binding residues.

It has been demonstrated that CTPR modules tend to self-assemble via head-to-tail interactions when terminal cysteines are present in crystal forms, solid films, and solution^{87,88}. The coordination of certain amino acids (histidines, cysteines, or tyrosines) with transition metals like copper, nickel, cobalt, or zinc is another method that is gaining popularity to provide protein-protein assembly. This method results in assemblies with different geometric structures based on the bonding between the amino acids that play the role of coordination residues and the metals^{89,90}. As the assemblies of arrayed proteins fall in nanometer distances, this biomolecular platform is very promising to organize multi-enzyme systems where intermediate transport between active site is somehow a rate-limiting step (*i.e.*, in presence of competing enzymes). Hence enzyme cascades that require NADH in situ recycling are excellent examples to assess the efficiency of these protein-based scaffolds in enhancing the cascade performance. Numerous methods relating to NADH cofactor recycling via enzyme assembly have been studied during the past ten years due to the high demand for this class of cofactors¹².

4.1.3. FDH/LAlaDH bi-enzyme cascade on CTPR proteins

The present work is based on the development of two protein-protein assemblies of two fusion TPR proteins formed by CTPRs and two dehydrogenase enzymes (SCAB-FDH and SCAB-AlaDH) by which new organized multi-enzyme systems are obtained. The bases of the assembly techniques were selective head-to-tail interactions, promoted by single engineered cysteine at the N- and C-terminal ends, resulting in the production of a disulfide covalent bonds⁸⁷, and metal-driven assembly, in which tailored histidines makes coordination bonds with copper (II) ion forming a metal-protein interaction that controls the protein-protein assembly⁹⁰ (Figure 4.3). To accomplish this objective, we have selectively assembled FDH and AlaDH by post-translationally fusing them to SCAB proteins. In this enzyme cascade, while AlaDH catalyzes the asymmetric reduction of α -ketoacids to

enantiopure α -amino acids at the expense of NADH, FDH that uses formic acid as an ancillary substrate to in situ recycle NADH. Through a variety of analytical tools we confirm that the synthetic CTPR-based scaffolds successfully assemble the orthogonal FDH-AlaDH bi-enzymatic cascade, paving the way for the assembly of enzymes on previously designed protein scaffolds to produce novel, complex, and organized multi-enzyme systems⁹⁴.

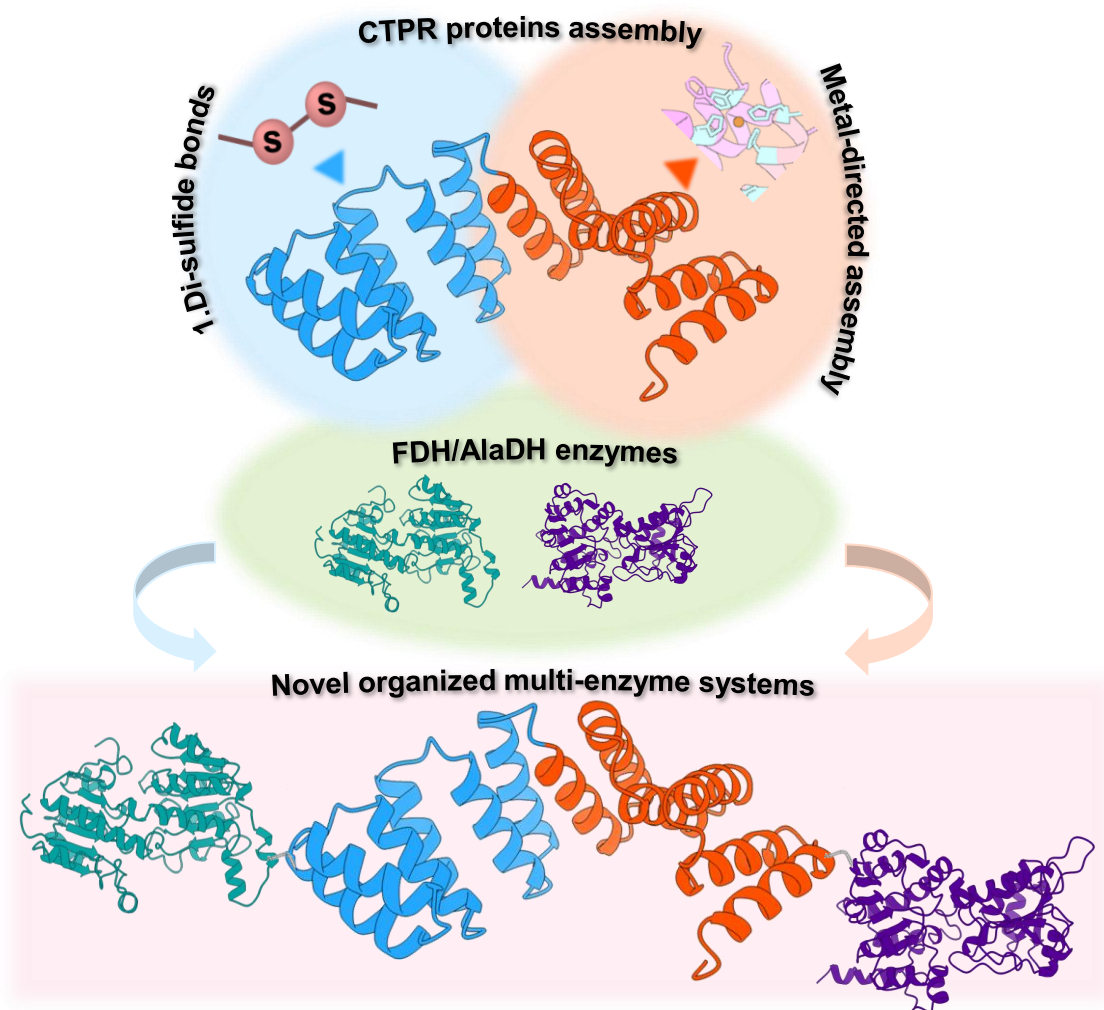


Figure 4.3. Schematic representation of the FDH/AlaDH bi-enzymatic cascade on CTPR proteins. These supramolecular assemblies take place based on head-to-tail interactions covalently anchor-bound interaction via di-sulfide bonds and anchor-bound interaction by coordination histidines and a metal ion; Cu^{2+} . CTPR modules are depicted in blue and red (PD ID: 2HYZ), along with FDH in turquoise, (PDB ID: 5DNA) and AlaDH in purple (PDB ID: 1PJB) enzymes.

4.2. Results and discussion

4.2.1. Design of the SCAffolding Bricks (SCABs)

In this work, we explored the potential of using genetically programmed intermolecular staples to bridge CTPR domains and assemble multi-enzyme systems in a spatially

organized manner. To achieve this, two strategies were employed to design orthogonal inter-molecular staples that could stabilize the assemblies formed through the intrinsic head-to-tail interactions between CTPR modules⁸⁷. Specifically, we used cysteine-mediated di-sulfide bonds and metal-mediated assembly as stapling chemistries to lock the assemblies.

First, inspired by our previous work in which head-to-tail interactions were stapled by selective di-sulfide bonds between unique N- and C-terminal cysteines within a CTPR20 (CTPR wild-type protein with 20 repeats) module, leading to the formation of ordered protein nanofibers⁸⁷, we aim to design orthogonal scaffolding units that would encode directional order within the assembly. Towards this aim, we exploited for the first time the two orthogonal interfaces of a CTPR protein (the intra- and inter-repeat packing interfaces) to design two unique orthogonal modules that can assemble in a directional manner, avoiding the self-polymerization that was described for the pseudo infinite linear fiber formation⁸⁷. Each CTPR domain is composed by two helices; A and B, that are genetically fused and interact with each other to form a A-B intra-repeat interface. Furthermore, CTPR proteins present a second inter-repeat packing interface composed mostly of the interaction between the helix B, and the helix A of the following repeat (Figure 1a). These two interfaces can be further employed to drive oligomerization between different CTPR proteins based on the crystal structure of a 8-repeat long CTPR protein (CTPR8, PDB ID: 2HYZ¹²⁶), we designed two novel CTPR-based modules with orthogonal unsatisfied interfaces (intra- and inter-), whose sequences are A-B-A-B-A for SCAB1 and B-A-B-A-B for SCAB2. To ensure stability, each module comprises two and a half repeats. Consecutive alternating SCAB1 and SCAB2 modules can assemble into linear arrays that recapitulate the CTPR super helical extended structure (Figure 4.4a). Next, we introduced a unique cysteine at the C- and N-terminal ends of the SCAB1 and SCAB2 modules to generate SCAB1_C and SCAB2_C, respectively, to promote inducible covalent stapling through the formation of a di-sulfide bond at the intra-repeated interface (Figure 4.4.b).

Following the same rationale, we designed a metal-driven assembly within CTPR modules, as previously used to drive complex protein assemblies⁹⁰. The formation of such a metal-driven staple, which brings together the two SCABs modules, relies on the strength and selectivity of metal-protein interactions⁸⁹. CTPR systems have already been engineered to display metal binding sites based on histidines and cysteines at non-conserved positions, which selectively coordinate metal ions for the subsequent stabilization of nanomaterials^{123,105,127}. For the design of two new SCABs modules, we introduced two metal-coordination bi-histidine at the C- or N-terminal of a SCAB module composed of three repeats (A-B-A-B-A-B sequence). Using previous design principles for metal coordination^{89,90}, we introduced four coordinating histidines (His) at positions 2, 6, 9, and 13 of the A helix in single CTPR domain (His module, H) (Figure 4.4c). The side chain conformations and backbone geometry of the His residues were modelled and found to be compatible with the CTPR structure. For scaffold stability purposes, modules comprising three repeats were generated by combining the newly engineered His module with wild-type (WT) CTPR units to obtain CTPR3 proteins with WWH and HWW modules named SCAB1_H and SCAB2_H, respectively.

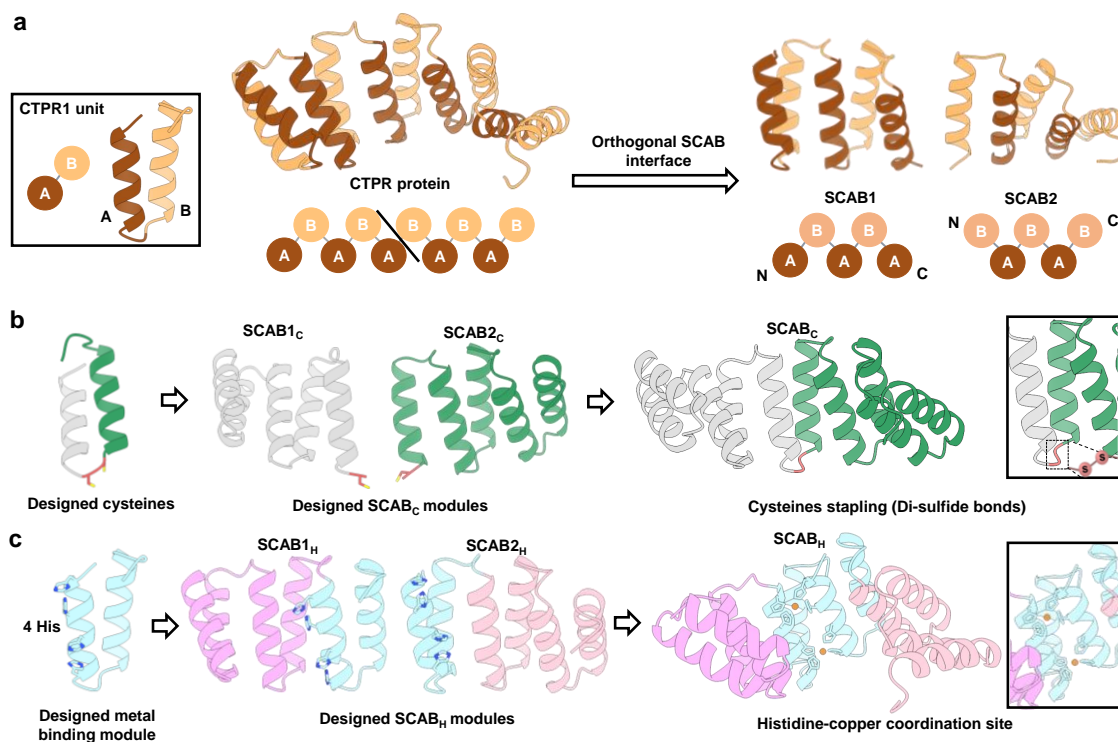


Figure 4.4. Design of two SCAffolding Bricks (SCAB) systems for their selective and orthogonal assembly through disulfide or metal coordination. a) A CTPR unit is comprised of two helices (A-B). CTPR proteins comprise of arrays of CTPR units can be splitted in two unsatisfied CTPR modules (SCAB1: A-B-A-B-A and SCAB2: B-A-B-A-B) with unsatisfied orthogonal interfaces (intra- and inter-) that will drive directional assembly based on intrinsic head-to-tail interactions. b) SCAB1_C and SCAB2_C modules in which two cysteines have been introduced at N- and C- terminal positions. Scheme of the assembly of SCAB1_C and SCAB2_C promoted by the formation of a di-sulfide bond (di-sulfide bond between the modules is shown in red), resulting in SCAB_C. c) The engineered His module (H) display four histidines for metal coordination at positions 2, 6, 9 and 13 positions (shown in cyan). SCAB1_H and SCAB2_H modules comprise three repeats, which combine two WT modules (pink and light pink) and a H module (cyan) at the C-terminal or N-terminal end repeat, respectively. Scheme of the assembly of SCAB1_H and SCAB2_H promoted by copper-driven interaction through the histidine residues, resulting in SCAB_H. The 3D structures of the SCAB modules have been modelled based on the crystal structure of the CTPR8-WT (PDB ID: 2HYZ) ¹²⁶.

The coordination sites that emerged upon the metal-directed interaction of the SCAB1_H and SCAB2_H modules were simulated by CheckMyMetal online server (Figure 4.5), and the metal-coordination distances were validated computationally with the aforementioned online server to ensure compatibility with the sites described for natural metal-coordinating proteins. The resulting intra-repeat interface comprises two tetra-histidine high affinity sites that will bond the two SCAB modules in presence of a divalent metal.

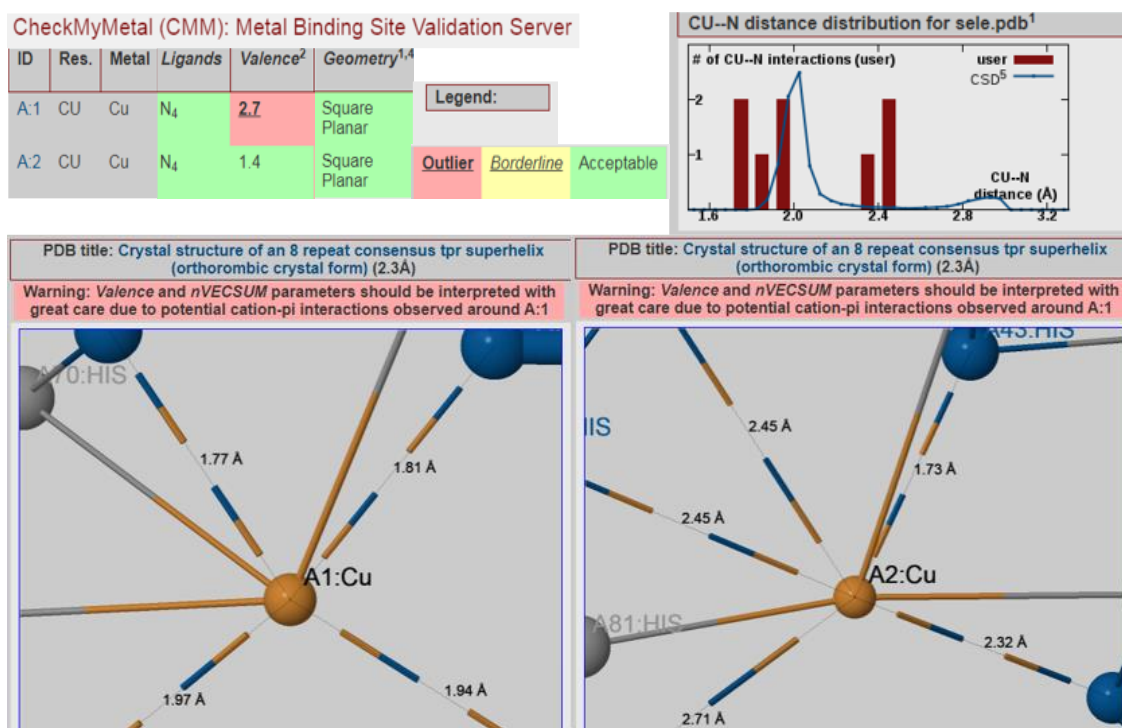
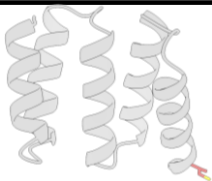
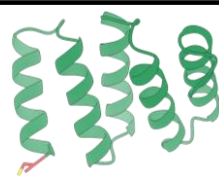

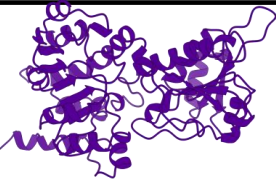
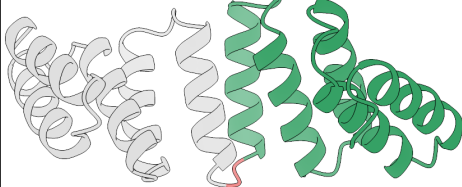
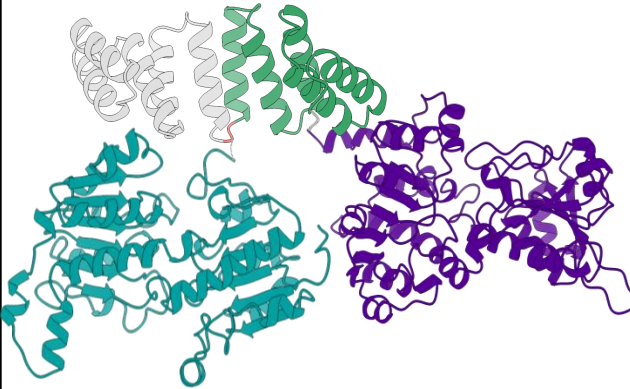
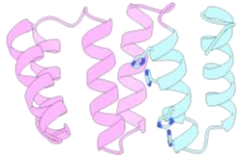
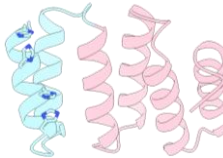
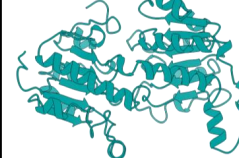
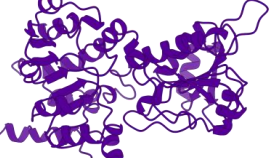
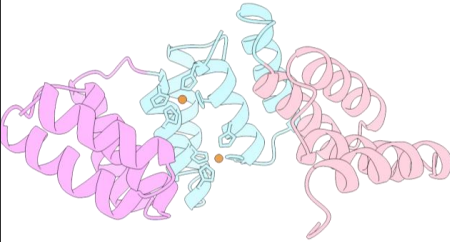
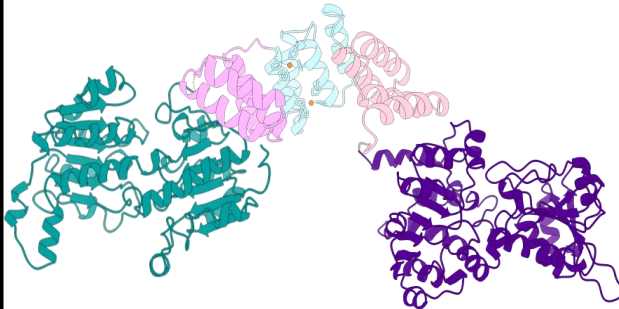


Figure 4.5. Validation of copper-histidine coordination by SCAB_H modules. The proposed two histidine-copper binding sites were validated by Metal Binding Validation Server: CheckMyMetal (CMM); *Check My Metal Server - Metal Sites (minorlab.org)*¹²⁸. Top: The designed coordination sites comprised of four coordinating histidines and copper metal were studied. The elemental composition of the coordination sphere (histidine ligands) and the valence of the coordinated copper (approximate valence; Cu²⁺) were evaluated. The theoretical square geometry promoted by histidine-copper bonding was shown. The color legend determined the validation of the parameters (green, acceptable; yellow, borderline; and red, outlier). Middle: Histogram of the distribution of Cu atoms coordinated with histidine ligands. Bottom: Graphical image of the two coordination sites between the histidines and the copper metal. The distance between histidines and copper ranges from 1.7 Å to 2.5 Å.

These newly designed sets of SCAB pair modules, SCAB_{1C} and SCAB_{2C}, and SCAB_{1H} and SCAB_{2H} (Tables 4.1 and 4.2) were expressed, purified, and characterized.

Table 4.1. Description of the protein elements involved in SCAB_C, SCAB_H, FDH/AlaDH@SCAB_C, and FDH/AlaDH@SCAB_H assemblies. The 3D model structures for the SCAB_C modules (SCAB_{1C}, gray and SCAB_{2C}, green) together with the designed cysteines at the N- and C-terminal ends for the assembly via reversible di-sulfide bonds are marked in red. The 3D model structures for the SCAB_H modules, SCAB_{1H}, and SCAB_{2H}, are shown. Each SCAB_H is composed of 3 repeat modules combining two WT CTPR modules in light or dark pink and an engineered metal coordinating module with 4 histidines (H module) for copper coordination (cyan). The 3D model structures of the SCAB modules have been modeled using the crystal structure of the WT CTPR8 (PDB ID: 2HYZ). The 3D model structures of the FDH (PDB ID: 5DNA) and AlaDH (PDB ID: 1PJB) enzymes in turquoise and purple, respectively.

SCAB module 1 (SCAB1 _C)	SCAB module 2 (SCAB2 _C)	Enzyme 1 (FDH)	Enzyme 2 (AlaDH)
			
Assembly by di-sulfide bonds (SCAB _C)		Assembled enzymes by di-sulfide bonds (FDH/AlaDH@SCAB _C)	
			
SCAB module 1 (SCAB1 _H)	SCAB module 2 (SCAB2 _H)	Enzyme 1 (FDH)	Enzyme 2 (AlaDH)
			
Assembly by metal-driven assembly (SCAB _H)		Assembled enzymes by metal-driven assembly (FDH/AlaDH@SCAB _H)	
			

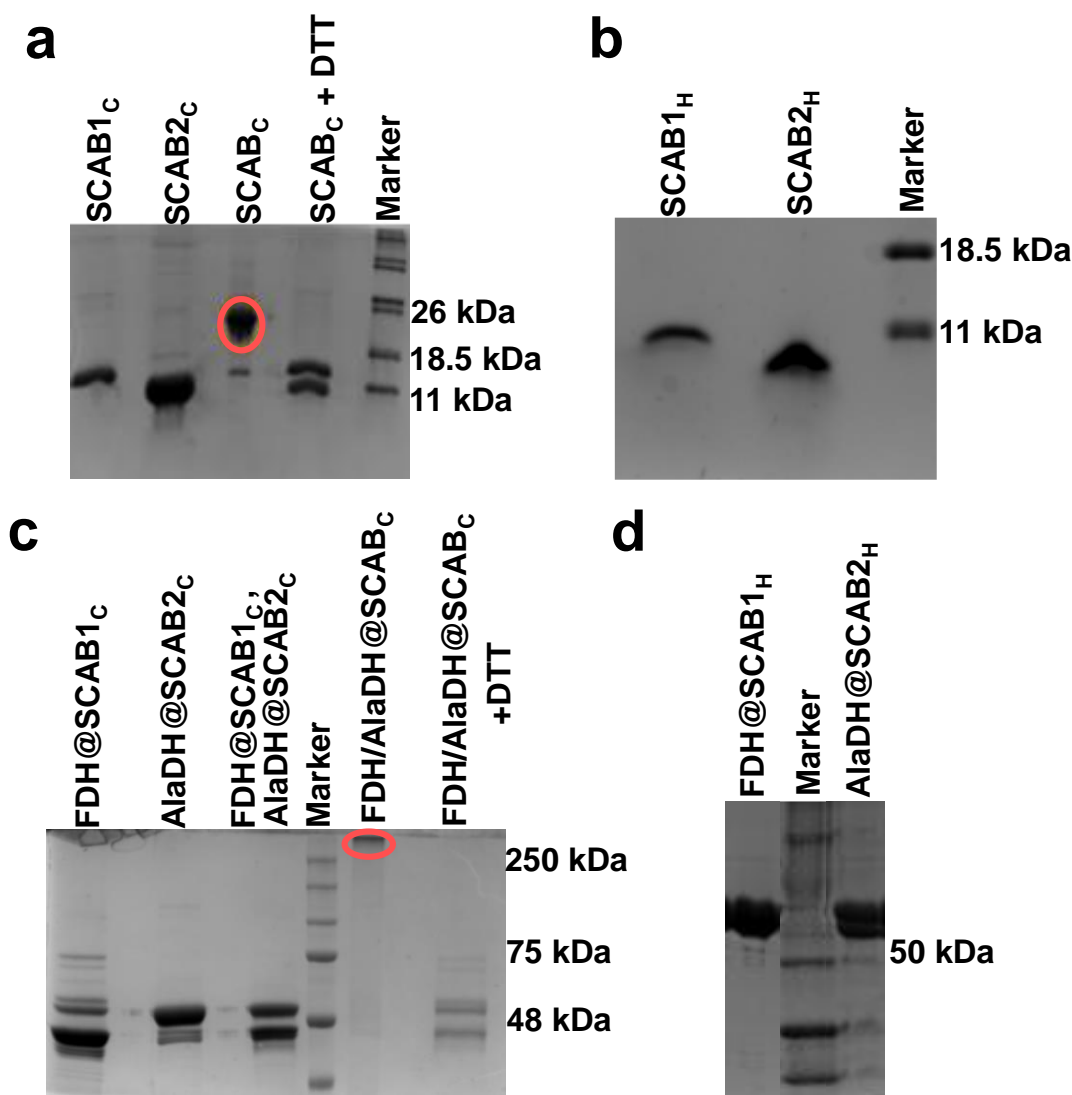


Figure 4.6. SDS-PAGE analysis of SCAB modules and SCAB-enzyme fusions. a) SDS-PAGE electrophoresis gel of purified SCAB_C modules and SCAB_C reversible assembly. From left to right: SCAB_{1c}, SCAB_{2c}, and SCAB_c without and in the presence of reducing agent, dithiothreitol (DTT). The SCAB_c assembly (circled in red) was previously purified by size exclusion chromatography (SEC). b) SDS-PAGE electrophoresis gel of purified SCAB_H modules: SCAB_{1h} and SCAB_{2h}. c) SDS-PAGE electrophoresis gel of purified SCAB_C-enzyme fusions and FDH/AlaDH@SCAB_C reversible assembly. From left to right: FDH@SCAB_{1c}, AlaDH@SCAB_{2c}, non-assembled FDH@SCAB_{1c}, AlaDH@SCAB_{2c} and FDH/AlaDH@SCAB_c without and in the presence of reducing agent, DTT. The FDH/AlaDH@SCAB_c assembly (circled in red) was previously purified by SEC. d) SDS-PAGE electrophoresis gel of purified SCAB_H-enzyme fusions: FDH@SCAB_{1h} and AlaDH@SCAB_{2h}.

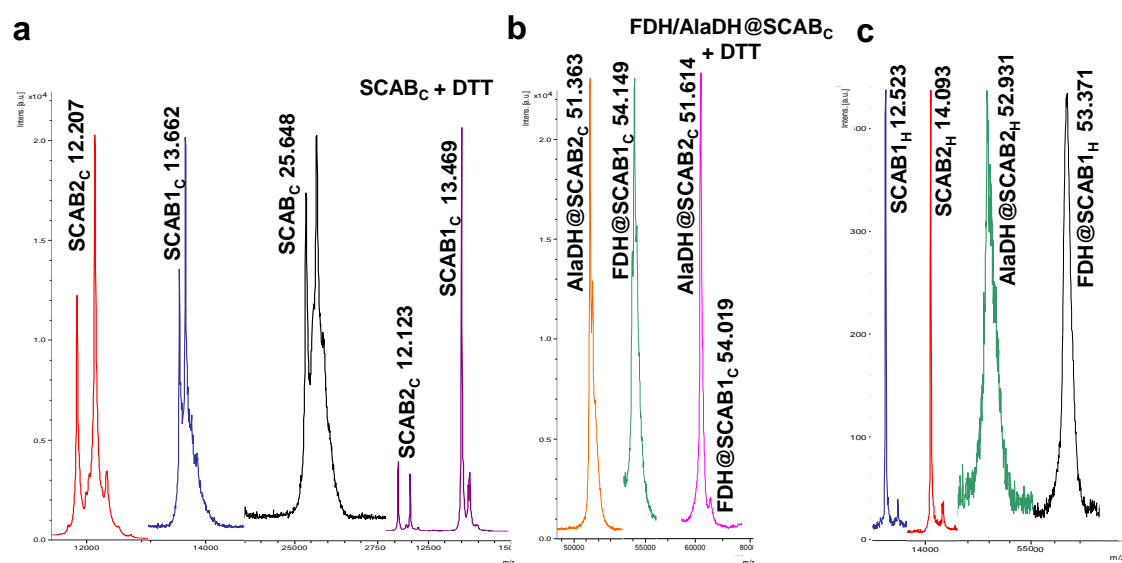


Figure 4.7. MALDI-TOF spectra and calculated masses for SCAB modules, SCAB-enzyme fusions, and assemblies via di-sulfide stapling. a) MALDI-TOF spectra and calculated masses from left to right: SCAB_{2c} (in red; peak at 12.207 kDa), SCAB_{1c} (in dark blue; peak at 13.662 kDa), SCAB_c assembly (in black; peak at 25.648 kDa) and SCAB_c assembly with the presence of DTT reducing agent resulting in SCAB_{1c} and SCAB_{2c} (in purple; two peaks at 13.469 kDa and 12.123 kDa). b) MALDI-TOF spectra and calculated masses from left to right: AlaDH@SCAB_{2c} (in orange; peak at 51.363 kDa), FDH@SCAB_{1c} (in green; peak at 54.149 kDa) and FDH/AlaDH@SCAB_c assembly with the presence of DTT reducing agent resulting in FDH@SCAB_{1c} and AlaDH@SCAB_{2c} (in pink; two peaks at 54.019 kDa and 51.614 kDa). Masses calculated by amino acid compositions resulting in 13.682 kDa for SCAB_{1c}, 12.664 kDa for SCAB_{2c}, 26.346 kDa for SCAB_c assembly, 54.073 kDa for FDH@ SCAB_{1c} and 51.342 kDa for AlaDH@SCAB_{2c}. c) MALDI-TOF spectra and calculated masses from left to right: SCAB_{1H} (in blue; peak at 12.525 kDa), SCAB_{2H} (in red; peak at 14.220 kDa), FDH@SCAB_{1H} (in black; peak at 53.371 kDa) and AlaDH@SCAB_{2c} (in green; peak at 52.931 kDa). Masses calculated by the amino acid compositions of SCAB_{1H}: 12.493 kDa, SCAB_{2H}: 14.188 kDa, FDH@SCAB_{1H}: 53.091 kDa and AlaDH@SCAB_{2H}: 52.032 kDa.

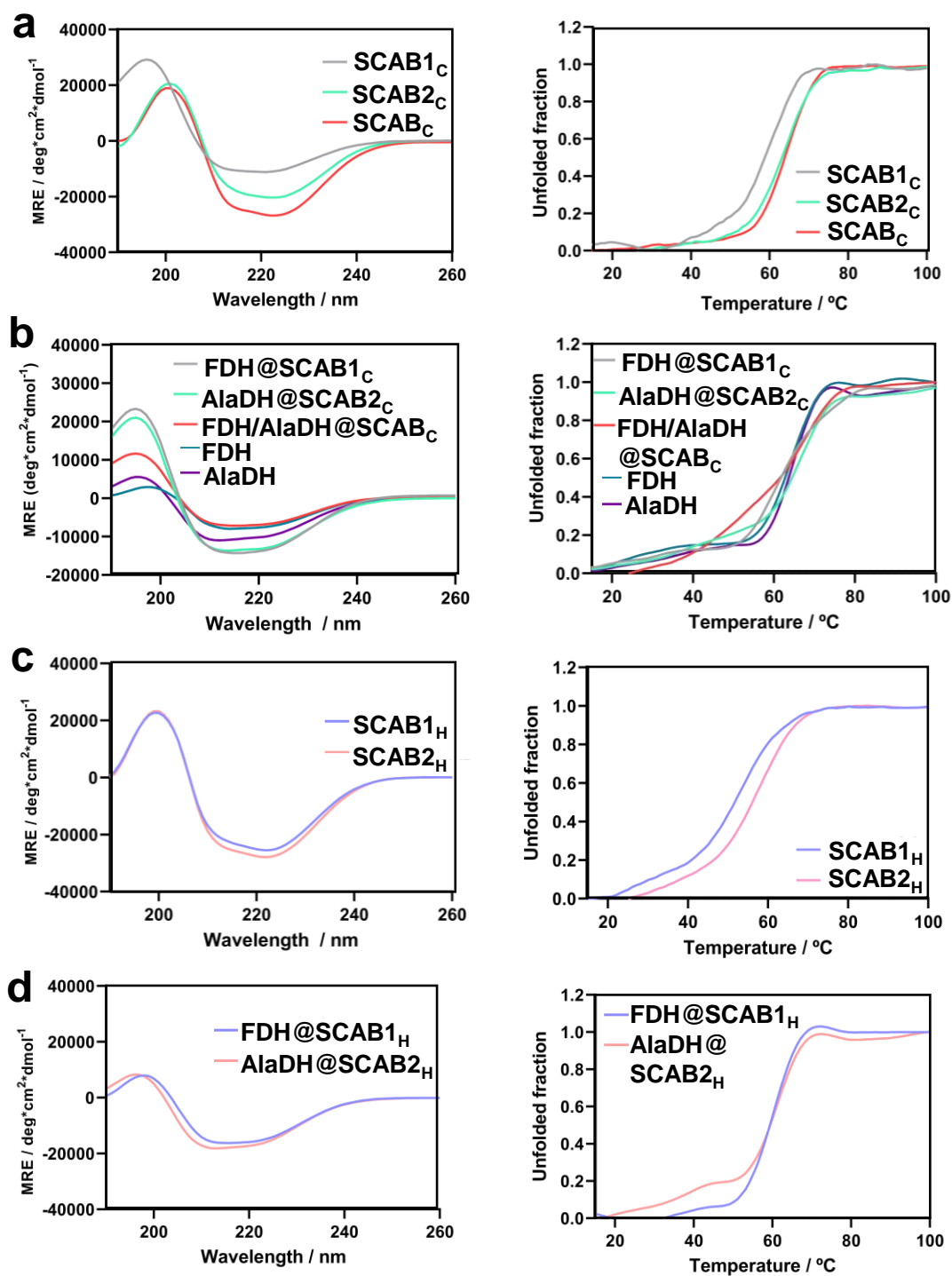


Figure 4.8. Circular dichroism spectra and thermal denaturation curves of SCAB modules and SCAB-enzyme fusions, FDH and AlaDH free enzymes, and assemblies via di-sulfide stapling. a) Left: circular dichroism spectra of SCAB_{1C} and SCAB_{2C} modules, and SCAB_C assembly. Right: thermal denaturation curves of SCAB_{1C} and SCAB_{2C} modules, and SCAB_C assembly. b) Left: circular dichroism spectra of FDH@SCAB_{1C} and AlaDH@SCAB_{2C} fusions, FDH/AlaDH@SCAB_C assembly, and FDH and AlaDH free enzymes. Right: thermal denaturation curves of FDH@SCAB_{1H} and AlaDH@SCAB_{2C} fusions, FDH/AlaDH@SCAB_C assembly, and FDH and AlaDH free

enzymes. SCAB1_C and FDH@SCAB1_C in grey, SCAB2_C and AlaDH@SCAB2_C in green, SCAB_C and FDH/AlaDH@SCAB_C assemblies in red, FDH enzyme in turquoise and AlaDH enzyme in purple. c) Left: circular dichroism spectra of SCAB1_H and SCAB2_H modules. Right: thermal denaturation curves of SCAB1_H and SCAB2_H modules. d) Left: circular dichroism spectra of FDH@SCAB1_H and AlaDH@SCAB2_H fusions. Right: thermal denaturation curves of FDH@SCAB1_H and AlaDH@SCAB2_C fusions. SCAB1_H and FDH@SCAB1_H in light purple and SCAB2_H and AlaDH@SCAB2_H in light pink.

Next the assembly of the two pairs of SCAB modules was evaluated through the two distinct assembly procedures.

The assembly of SCAB_C was performed stepwise in solid phase using a Ni-NTA affinity resin to enable a multi-modular assembly, avoiding non-specific polymerization of modules with two cysteines. First, SCAB1_C was immobilized on the resin through its His-tag, and then an excess of SCAB2_C without His-tag was added and incubated overnight at 37 °C. After washing the excess of SCAB2_C, the final assembly was eluted and analyzed (Figure 4.9a). Electrophoresis analysis under non-reducing conditions showed the presence of SCAB_C dimer and the disassembly after reduction of the di-sulfide bond with dithiothreitol (DTT) (Figure 4.9a). MALDI-TOF spectra demonstrate that the size of the SCAB_C assembly, determined by spectra is 25.7 kDa, in agreement with the expected mass of the dimer, and the stoichiometry of the assembly is 1:1 (SCAB1_C:SCAB2_C), thus validating the proposed scaffolding strategy. After addition of DTT, we demonstrated the reversibility of assembly as both SCAB1_C and SCAB2_C monomers (13.5 kDa and 12.1 kDa) were also detected by mass spectrometry (Figure 4.6a and 4.7a). Circular dichroism analysis revealed that the α -helical secondary structure of the SCAB_C is preserved within the assembly, and the system presents similar stability to the individual components as determined by the thermal melting temperature (T_M SCAB1_C 59.03 °C and T_M SCAB2_C 63.03 °C) (Figure 4.8a). Finally, the SCAB_C assembly was analyzed by SEC. As expected, this analysis revealed a significant increase in the elution peak volume for SCAB_C assembly (11.15 mL) compared to SCAB_C modules separately (12.89 mL and 13.09 mL, specifically for SCAB1_C and SCAB2_C), illustrating the orthogonal assembly of the SCAB_C modules (Figure 4.9b).

The metal-driven assembly of SCAB1_H and SCAB2_H was carried out in solution using different metal salts, CuSO₄, NiSO₄ and CoSO₄. This assembly was not made in solid phase due to the presence of histidines, since the affinity of these for the Ni-NTA resin would not allow the formation of the assembly. Therefore, SCAB1_H and SCAB2_H modules were incubated in solution at a protein concentration of 50 μ M for 1 h at 50 °C with two equivalents of the metal salt per protein (Figure 4.9c).

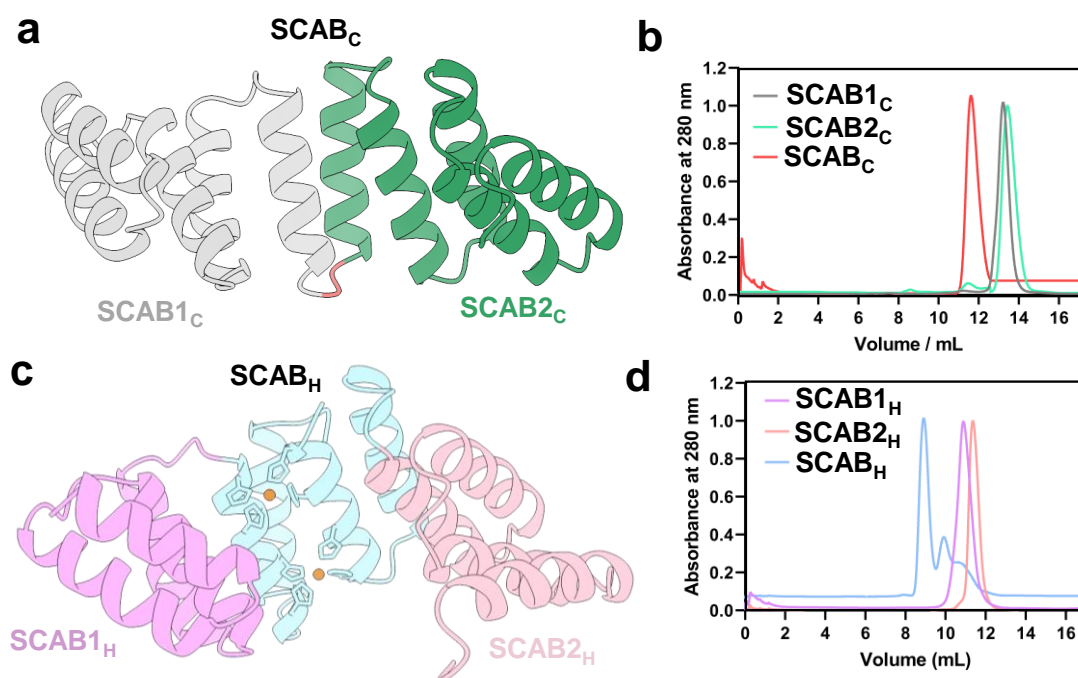


Figure 4.9. Assembly SCAffolding Bricks SCAB_C and SCAB_H. a) Schematic representation of SCAB_C assembly-based scaffold composed by SCAB1_C (in grey) and SCAB2_C (in green), resulting in the SCAB_C assembled system. Introduction of the assembly concept. The assembly was based on head-to-tail-interactions between complementary modules and a di-sulfide bond staple. b) SEC chromatograms through a Superdex S75 (10/300) GL column monitored at 280 nm wavelength for the SCAB_C modules (SCAB1_C and SCAB2_C), and SCAB_C assembled system. c) Schematic representation of the SCAB_H assembly-based scaffold composed by SCAB1_H (in bright pink and His module in cyan) and SCAB2_H (in light pink and His module in cyan), resulting in the SCAB_H assembled system. The assembly was based specific metal coordination between SCAB1_H and SCAB2_H. d) SEC chromatograms through a Superdex S75 (10/300) GL column monitored at 280 nm wavelength by SEC for the SCAB_H modules (SCAB1_H and SCAB2_H), and SCAB_H assembled system.

The size of the assembly in the presence of the different metal ions was analyzed by SEC, showing that the largest shift in size occurred in the presence of the Cu²⁺ metal ion (Figure 4.10).

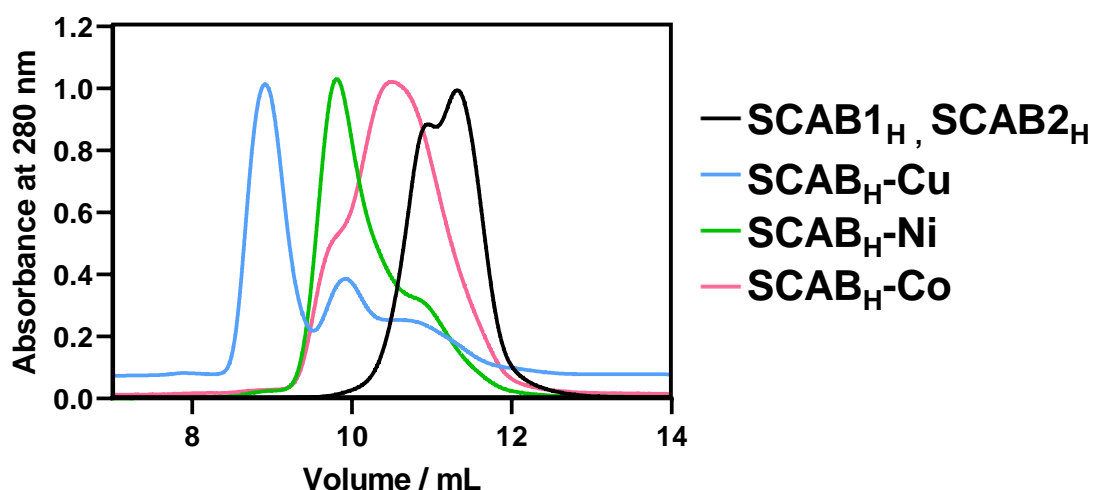


Figure 4.10. Purification of SCAB_H assembled systems by SEC using a Superdex 75 (10/300) GL column. Chromatogram monitored at 280 nm wavelength by SEC for the mix of SCAB_H modules (SCAB1_H and SCAB2_H) without metal ions, and SCAB_H assembled systems with different metal ions: Cu²⁺, Ni²⁺, and Co²⁺.

Therefore, copper was selected for further optimization of the assembly conditions in solution. For this purpose, a constant concentration of SCAB2_H was incubated with different concentrations of SCAB1_H in the presence of two equivalents of Cu as coordination metal for 1 h at 50 °C (Figure 4.11).

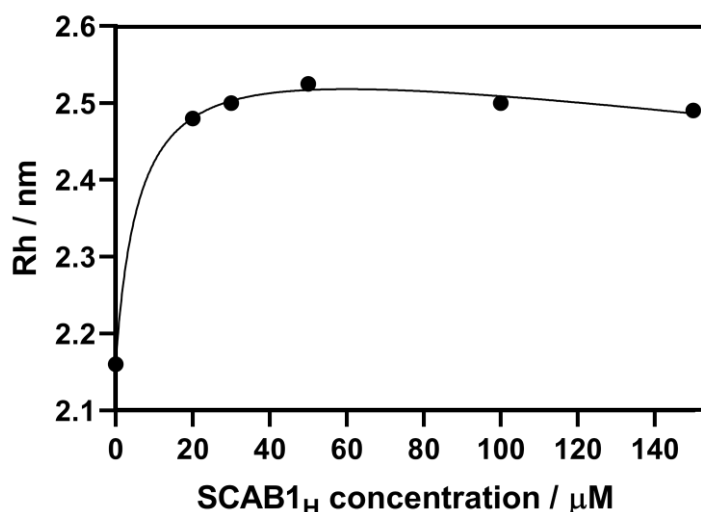


Figure 4.11. Assembly of SCAB1_H and SCAB2_H by monitoring the average apparent Rh in nm when SCAB2_H is titrated with increasing SCAB1_H concentrations. The apparent average Rh values were plotted for the different assembled SCAB_H systems and were measured by microfluidic diffusional sizing technology using Fluidity One W. In all SCAB_H assembled systems, SCAB2_H was labeled with Alexa Fluor-647 (AF647) dye. The different molar ratios of the elements composing the SCAB_H assembly: 0.4:1, 0.6:1, 1:1, 1:1, 1:2, and 1:3 for SCAB1_H:SCAB2_H, respectively. The average Rh increases in all the

analyzed ratios of SCAB1_H:SCAB2_H, with the highest increase being seen at the 1:1 molar ratio. The optimal conditions of the SCAB_H assembly were at 50 μ M protein concentration and 1:1 molar ratio for the SCAB_H modules. After optimizing the process, the SCAB_H assembly was purified by SEC, which showed a clear shift in the elution peak volume (8.92 mL) when compared with the individual SCAB_H modules (10.87 mL and 11.37 mL, respectively for SCAB1_H and SCAB2_H), indicating the orthogonal assembly of the SCAB_H modules (Figure 4.9d). Finally, by inductively coupled plasma mass spectrometry (ICP-MS) we quantified 1.88 ± 0.94 Cu atoms per SCAB_H assembly, in accordance with the two metal-binding sites introduced in each SCAB_H.

4.2.2. Assembly, catalytic activity, and biosynthesis of L-Alanine of the multi-enzyme complexes through engineered SCAB Bio-bricks

Once we confirmed the effective orthogonal assembly of SCAB_C and SCAB_H bricks, we exploited them as scaffolding platforms to form multi-enzyme systems arrays with nanometric precision. To this end, we selected a two-enzyme system comprised of NADH dependent L-Alanine dehydrogenase from *Bacillus stearothermophilus* (AlaDH) and NADH cofactor regenerating formate dehydrogenase from *Candida boidinii* (FDH) that perform the biosynthesis of L-Alanine with *in-situ* NADH recycling. For the assembly of the bi-enzymatic systems, the previously validated SCAB modules were fused to the corresponding enzymes. AlaDH was fused at the C-terminus to SCAB2_C and SCAB2_H, whereas FDH was fused at the C-terminus to SCAB1_C and at the N-terminus to SCAB1_H, (since FDH fusion at the C-terminus of SCAB1_H interfered with the metal coordination). *In silico* predictions of the inter enzyme distance based on the 3D model of the inter-repeated stapled CTPRs domains suggest that these two enzymes were separated by 3.2 nm and 5.3 nm when assembled into the SCAB_C (Figure 4.12a) and SCAB_H (Figure 4.12c), respectively. Thus, we compared the assembly efficiency of the two-scaffold architectures to further study their effect on the performance of the model enzyme cascade mentioned above. The resulting enzymes variants fused with the SCAB modules were named as AlaDH@SCAB2_C, AlaDH@SCAB2_H, FDH@SCAB1_C, and FDH@SCAB1_H (Tables 4.1 and 4.2). These SCAB-enzyme fusions were purified, and their purity and masses were verified by SDS-PAGE gel (Figures 4.6c and 4.6d) and mass spectrometry (Figure 4.7b and 4.7c). The observed masses matched with the calculated mass from their corresponding amino acid sequence. Furthermore, circular dichroism analysis revealed that the secondary structure of the SCAB and the enzymes was negligibly affected by their direct fusion compared to individual SCAB modules and enzymes, which were also stable and showed similar thermal denaturation transitions (Figures 4.8c and 4.8d).

Once the SCAB-enzyme fusions were characterized, their assembly was carried out following the two strategies optimized above for the SCAB modules. The assembly of FDH/AlaDH@SCAB_C was performed as described for SCAB_C, with the only difference being that the SCAB_C-enzyme fusions (FDH@SCAB1_C and AlaDH@SCAB2_C) were first incubated in solution at 50 μ M protein concentration at 4 °C overnight and then immobilized on a Ni-NTA column through the His-tag FDH@SCAB1_C (Figure 4.12a). This assembly was performed in solution, as the formation of the disulfide bond did not occur in the solid phase probably due to the large size of the fusion proteins. Additionally, a lower assembly temperature was implemented to preserve enzyme activity and slow down the

assembly process of these larger proteins. The FDH/AlaDH@SCAB_C assembly was then eluted with imidazole and further purified by SEC. This analysis showed a notable increase in the elution peak volume of the assembly, which eluted at the exclusion volume (≈ 8 mL), indicating a large MW assembly when compared to the SCAB_C-enzyme fusions independently (10.81 mL and 13.11 mL, for FDH@SCAB1_C and AlaDH@SCAB2_C, respectively) (Figure 4.12b). The assembly was analyzed by SDS-PAGE gel and MALDI-TOF mass spectrometry (Figures 4.6c and 4.7b). Gel electrophoresis confirmed the FDH/AlaDH@SCAB_C assembly and the disassembly upon the reduction of the di-sulfide bond with DTT (Figure 4.6c). MALDI-TOF spectra for FDH/AlaDH@SCAB_C assembly did not show a peak corresponding to the assembly, probably due to its large molecular weight, as illustrated by SEC and SDS-PAGE analysis. However, the FDH@SCAB1_C and AlaDH@SCAB2_C monomers could be detected after addition of DTT to the assembled fraction (Figures 4.7b), which confirms the disulfide stapling of the SCAB_C-based scaffolds, enabling the complexation of the multi-enzyme system. Circular dichroism spectra showed that the structure of the FDH/AlaDH@SCAB_C was not significantly affected when compared with the individual elements (Figure 4.6b, left), thus the scaffolds affect the oligomerization state of the enzyme complex but not its secondary protein structure. The thermal denaturation curves show that all the individual elements as well as the assembly are stable and well folded since they show a cooperative denaturation curve (Figure 4.6b, right).

In parallel, we conducted the metal-driven assembly of FDH and AlaDH using the SCAB_H pair described above for the standalone SCAB_H modules (50 μ M protein concentration of each SCAB_H-enzyme fusion and 2 equivalents of CuSO₄ per protein concentration). Unfortunately, the assembly of SCAB_H needs high temperatures to occur that are incompatible with the enzymes. As the temperature of 4 °C is too low to promote this assembly, we selected 30 °C as a compromise temperature that allows the histidine-copper coordination without adversely affecting the activity of the assembled enzymes (Figure 4.12c). Under these conditions, we obtained FDH/AlaDH@SCAB_H metal-driven assembly, which was purified by SEC. The SEC-purified FDH/AlaDH@SCAB_H assembly showed a significant change in elution peak volume (8.55 mL) compared to the individual SCAB_H-enzyme fusions (10.89 mL and 9.42 mL, for FDH@SCAB1_H and AlaDH@SCAB2_H, respectively), demonstrating orthogonal assembly of the SCAB_H-enzyme fusions (Figure 4.12d). In addition, ICP-MS analysis confirmed the expected copper:SCAB_H stoichiometry with 1.98 ± 0.24 Cu atoms per FDH/AlaDH@SCAB_H assembly.

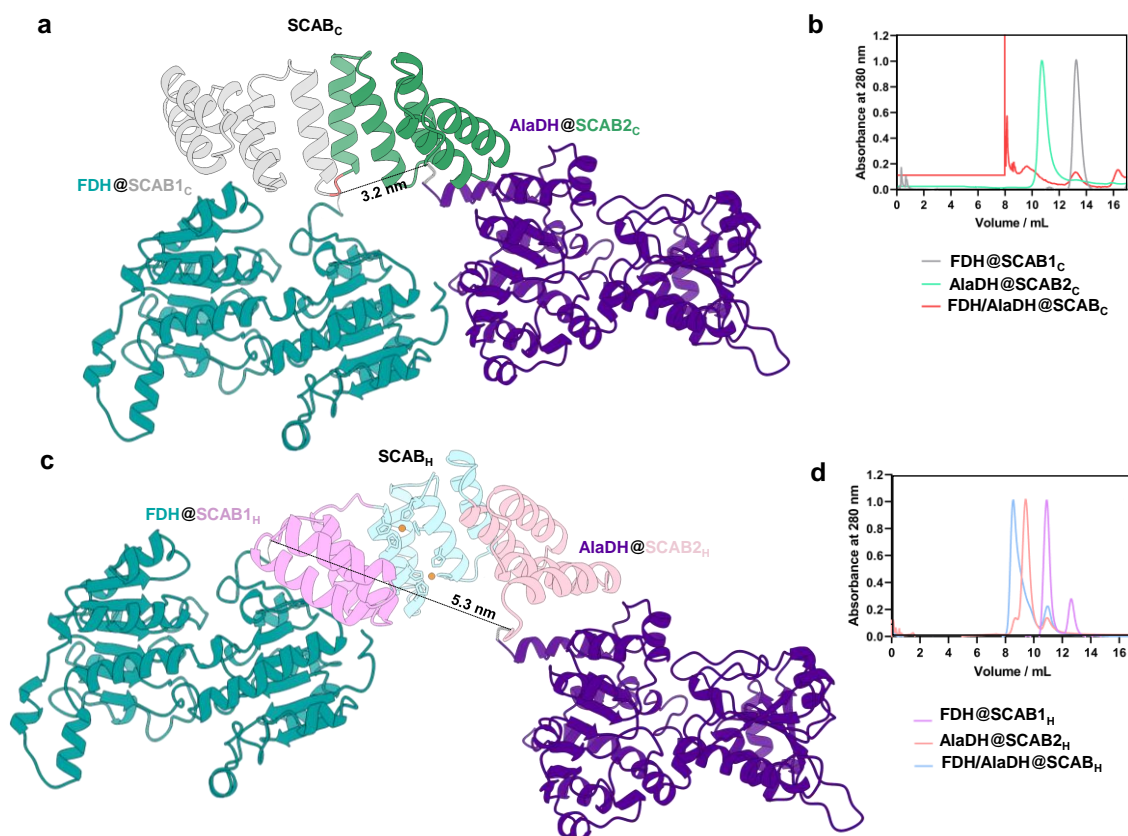


Figure 4.12. Schematic representation and characterization of FDH/AlaDH@SCAB_C and FDH/AlaDH@SCAB_H assemblies. a) Schematic representation of SCAB_C assembly composed of SCAB1_C (in grey) and SCAB2_C (in green), and their corresponding fused enzymes in their monomeric structural state, FDH (in turquoise, PDB ID: 5DNA), and AlaDH (in purple, PDB ID: 1PJB). The resulting FDH/AlaDH@SCAB_C assembled system, based on head-to-tail-interactions, is shown, and the distance between the enzyme anchoring points on SCAB_C scaffolds is approx. 3.2 nm. b) SEC chromatograms through a Superdex S200 (10/300) GL column monitored at 280 nm wavelength by SEC for the SCAB_C-enzyme fusions (FDH@SCAB1_C and AlaDH@SCAB2_C), and the assembled FDH/AlaDH@SCAB_C system. c) Schematic representation of SCAB_H assembly composed of SCAB1_H (WT modules in bright pink, and His module in cyan) and SCAB2_H (WT modules in light pink, and His module in cyan), and their corresponding fused enzymes in their monomeric structural state, FDH (in turquoise, PDB ID: 5DNA), and AlaDH (in purple, PDB ID: 1PJB). The resulting FDH/AlaDH@SCAB_H assembled system, based on metal-driven assembly, is shown, and the distance between the enzyme anchoring points on SCAB_C scaffolds is approx. 5.3 nm. d) SEC chromatograms through a Superdex S200 (10/300) GL column monitored at 280 nm wavelength for the SCAB_H-enzyme fusions (FDH@SCAB1_H and AlaDH@SCAB2_H), and the assembled FDH/AlaDH@SCAB_H system.

Next, we evaluated the effect of the assembly on the activity of the scaffolded enzymes. To that aim, we on UV-Vis assays and we determined measured the enzymatic activity of both free and assembled enzymes (Figure 4.13) using UV-Vis based assays (see Materials and Methods) and determined the Michaelis-Menten curves (Figure 4.14) and parameters (Figure 4.15).

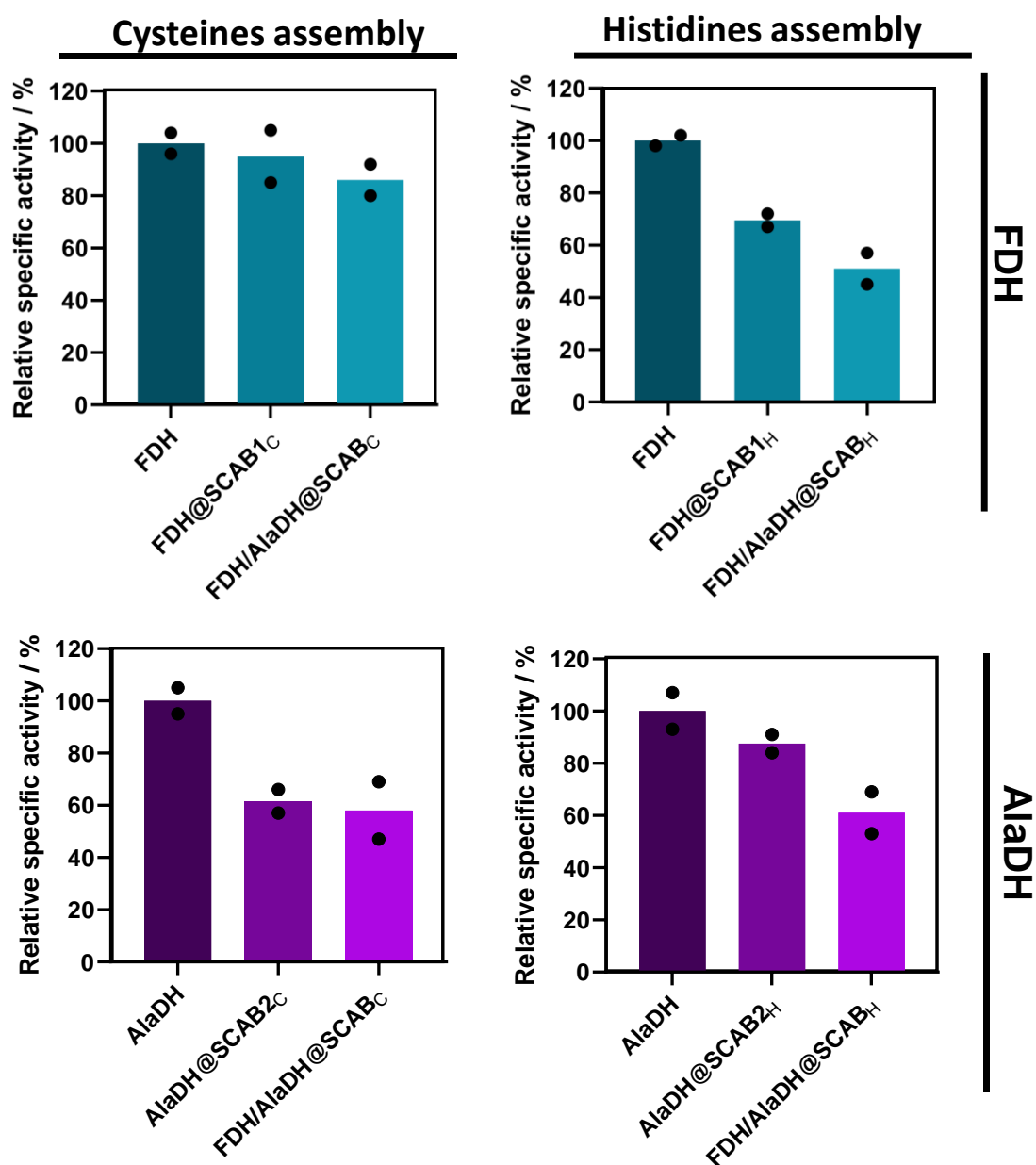
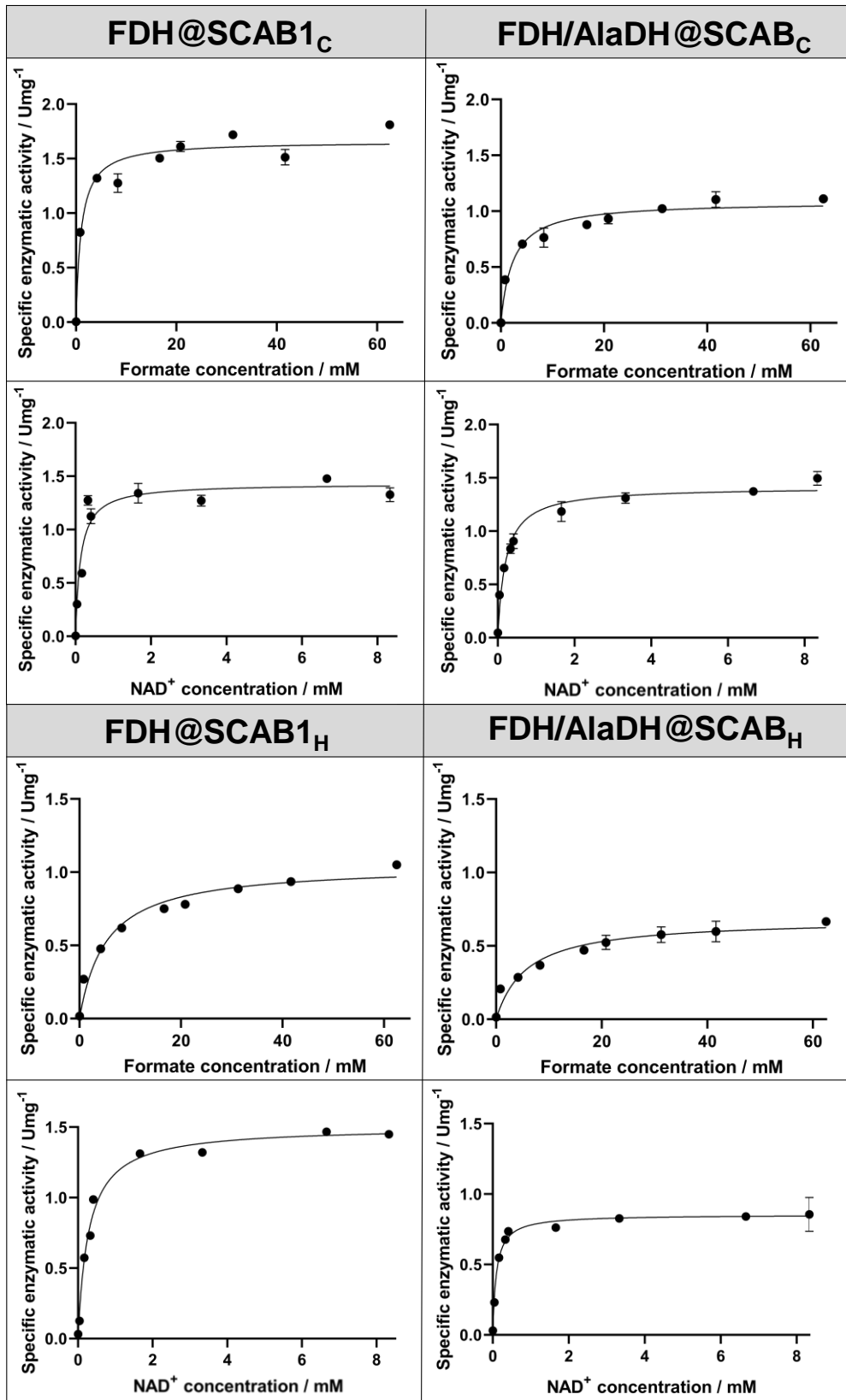


Figure 4.13. Relative activity of SCAB-enzyme fusions free and scaffolded on SCAB assemblies. Relative activity of free enzymes (FDH and AlaDH), SCAB-enzyme fusions free (FDH@SCAB1_C, FDH@SCAB1_H, AlaDH@SCAB2_C, AlaDH@SCAB2_H) and SCAB-enzyme fusions scaffolded (FDH/AlaDH@SCAB_C and FDH/AlaDH@SCAB_H). 100 % of AlaDH and FDH activities correspond to 22.1 U·mg⁻¹ and 0.83 U·mg⁻¹, respectively. The data are presented as the mean of two replicate experiments (n = 2).

a



b

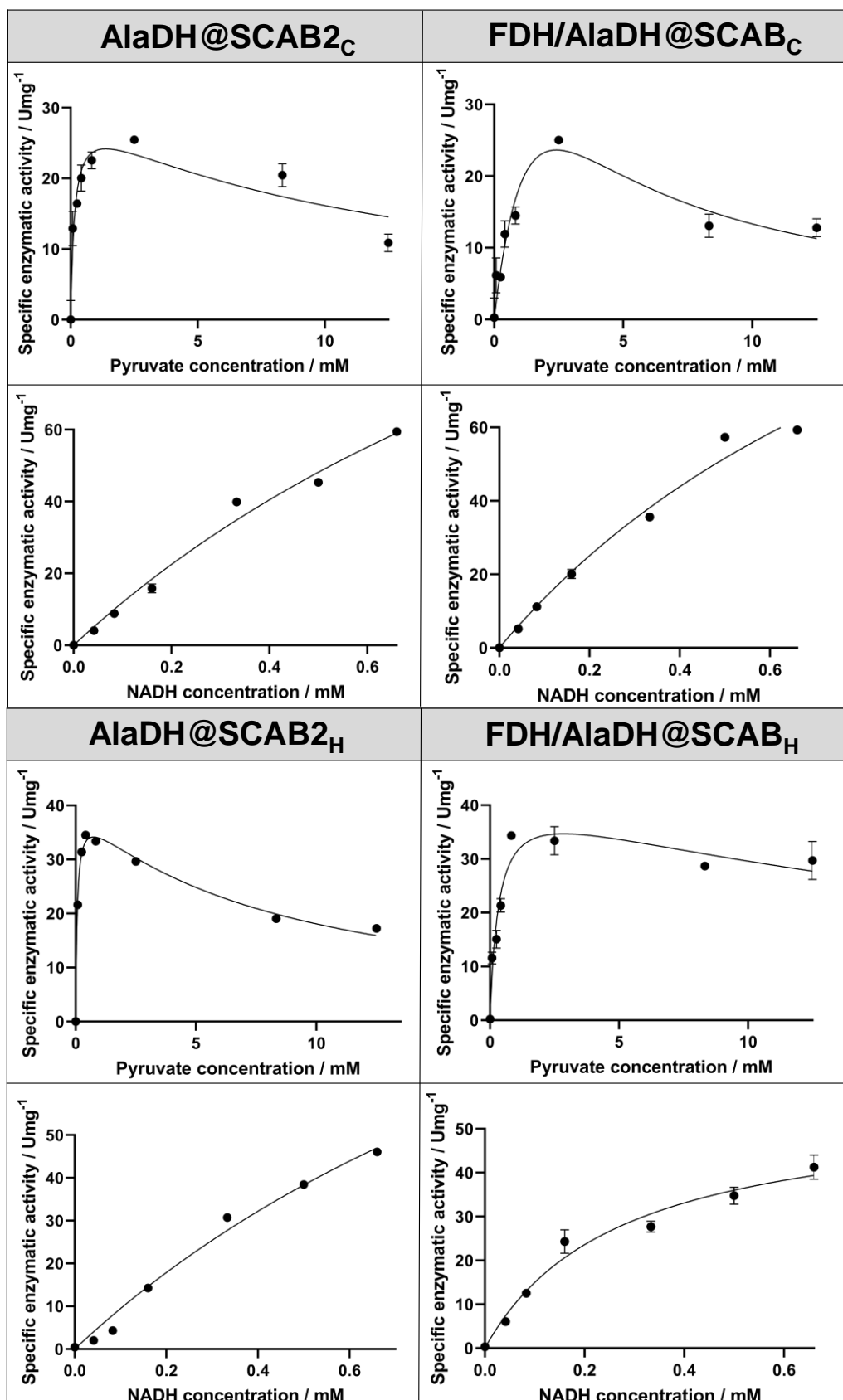


Figure 4.14. Michaelis-Menten curves for SCAB-enzyme fusions free and scaffolded.

a) Michaelis-Menten curves of SCAB-enzyme fusions free (FDH@SCAB_{1C} and FDH@SCAB_{1H}) and SCAB-enzyme fusions scaffolded (FDH/AlaDH@SCAB_C and FDH/AlaDH@SCAB_H) using different substrates: variable formate concentration at 1 mM NAD⁺ in 25 mM sodium phosphate buffer pH 7; and variable NAD⁺ concentration at 100 mM formate in 25 mM sodium phosphate buffer pH 7. b) Michaelis-Menten curves of SCAB-enzyme fusions free (AlaDH@SCAB_{2C} and AlaDH@SCAB_{2H}) and SCAB-enzyme fusions scaffolded (FDH/AlaDH@SCAB_C and FDH/AlaDH@SCAB_H) using different substrates: variable pyruvate concentration at 500 mM ammonium chloride, 0.5 mM NADH in 25 mM potassium phosphate buffer pH 8; and variable NADH concentration at 75 mM pyruvate, 500 mM ammonium chloride in 25 mM potassium phosphate buffer pH 8. Pyruvate kinetics were adjusted to a substrate inhibition model. The data are presented as the mean of three replicate experiments (n = 3), and error bars represent standard deviations.

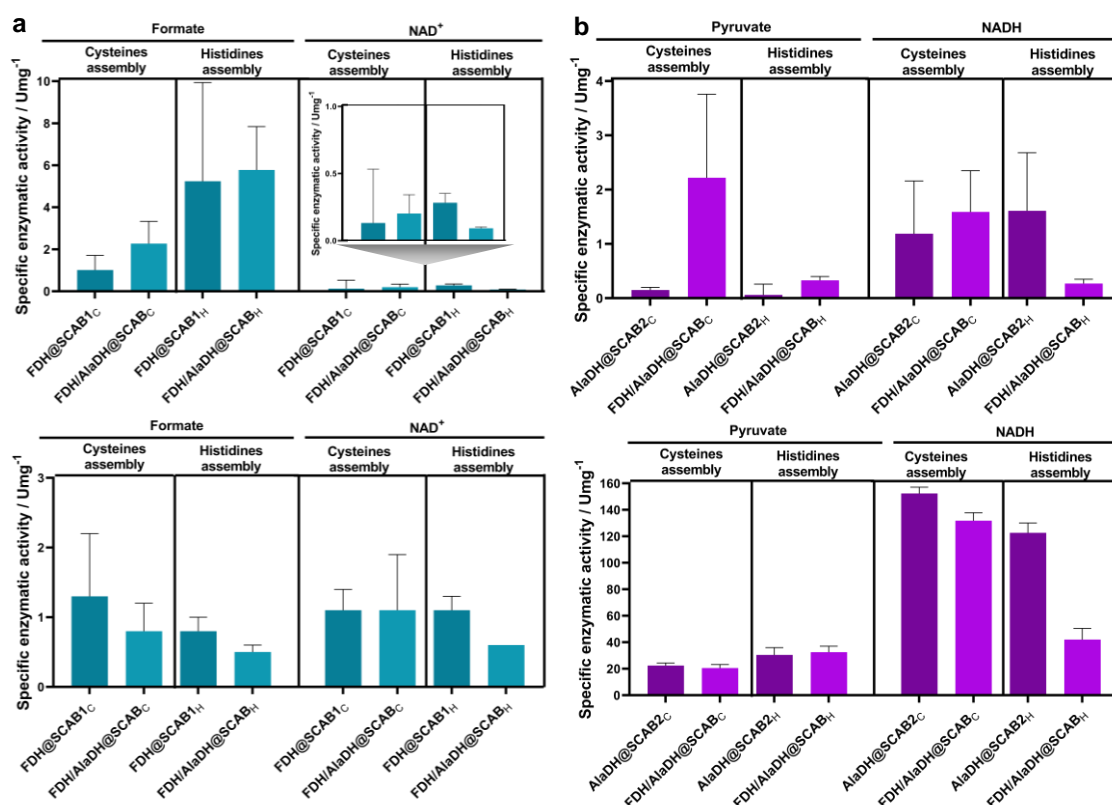


Figure 4.15. Michaelis-Menten steady-state kinetic parameters (K_M and k_{cat}) of free and scaffolded enzymes. a) K_M (top graphic) and k_{cat} (bottom graphic) data for FDH in the free SCAB-enzyme fusions, FDH@SCAB_{1C} and FDH@SCAB_{1H}, and in the scaffolded systems, FDH/AlaDH@SCAB_C and FDH/AlaDH@SCAB_H, in response to formate and NAD⁺ as distinct substrates. b) K_M (top graphic) and k_{cat} (bottom graphic) data for AlaDH in SCAB-enzyme fusions free, AlaDH@SCAB_{2C} and in AlaDH@SCAB_{2H}, and SCAB-enzyme fusions scaffolded, FDH/AlaDH@SCAB_C and FDH/AlaDH@SCAB_H, in response to pyruvate and NADH as distinct substrates.

The assembly of FDH on SCAB_H led decreased its activity to a higher extent than the assembly on SCAB_C. The lower activity of the FDH is mainly explained by the negative effect SCAB_H exerts over the function of FDH. However, AlaDH exhibits similar specific activity when assembled on both SCABs due to different reasons. Whereas its fusion to SCAB_C explains the lower activity of AlaDH mainly due to the assembly itself when forming the SCAB_H explains the lower activity of AlaDH on SCAB_C. The most deleterious scaffolding for the functionality of the two enzymes was SCAB_H where the k_{cat} of scaffolded FDH was significantly lower than its free counterpart and the K_M of the AlaDH towards pyruvate was much higher than its free counterpart.

Finally, we assessed the performance of the biosynthetic cascade when the enzymes were either scaffolded or free using 1 equivalent of pyruvate, 1.3 equivalents of sodium formate, 6.6 equivalents of ammonium chloride and substoichiometric amounts of NADH (150 times less than pyruvate) to synthesize L-Alanine (Figure 4.16a). Figure 4.16b shows that the scaffolded FDH/AlaDH@SCAB_C multi-enzyme system is 3.6 times faster than its free counterpart, achieving a specific productivity of $1.05 \text{ g} \times \text{g}_{\text{enzyme}}^{-1} \times \text{h}^{-1}$ and a titer of 9 mM of L-Alanine after 24 hours, corresponding to a chromatographic yield (CY) of 12 %. Meanwhile, under the same conditions and incubation period, the non-assembled system (free FDH@SCAB_{1C} and AlaDH@SCAB_{2C}) achieves 2.6-times lower CY. Contrastingly, the scaffolded FDH/AlaDH@SCAB_H multi-enzyme system displays the same catalytic performance as the free one, attaining a specific productivity of $0.17 \text{ g} \times \text{g}_{\text{enzyme}}^{-1} \times \text{h}^{-1}$ and a titer of 3.23 mM of L-Alanine after 24 hours, yielding 4.3 % CY. Under the same conditions, free FDH@SCAB_{1H} and AlaDH@SCAB_{2H} produced a CY of just 4.6 % (Figure 4.16c).

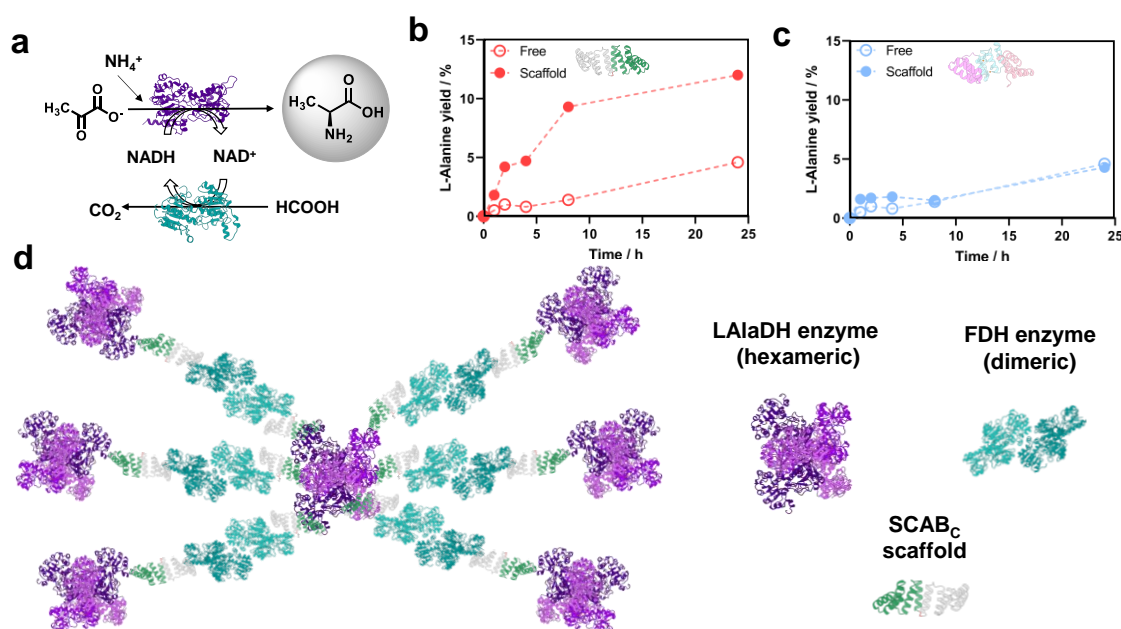


Figure 4.16. Catalytic performance of SCAB scaffolded enzymes for L-Alanine synthesis in batch-mode. a) Scheme of the catalytic cycle that was tested by HPLC. b) L-Alanine yield at 1:1 enzyme ratio of FDH:AlaDH for the free SCAB_C-enzymes

(FDH@SCAB1_C and AlaDH@SCAB2_C) and for the scaffolded SCAB_C-enzymes (FDH/AlaDH@SCAB_C). c) L-Alanine yield at 1:1 enzyme ratio of FDH:AlaDH for the free SCAB_H-enzymes (FDH@SCAB1_H and AlaDH@SCAB2_H) and for the scaffolded SCAB_H-enzymes (FDH/AlaDH@SCAB_H). Reaction mixture: 100 mM formate, 75 mM pyruvate, 500 mM ammonium chloride, and 0.5 mM NADH. d) Schematic representation of potential supramolecular assembly of the scaffolded enzymes on SCAB_C protein scaffolds driven the oligomeric state of AlaDH.

The volumetric productivity and the yield determined for the L-Alanine biosynthesis indicate that, only the spatial arrangement of the multi-enzyme system in SCAB_C exerts a positive effect on the overall throughput of the enzyme cascade. Given that FDH is the rate-limiting enzyme in this cascade (Figure 4.15), we hypothesize that the transport of NAD⁺/NADH between AlaDH and FDH when they are covalently scaffolded through SCAB_C is more efficient than when they are assembled into SCAB_H, the NADH recycling is more efficient scaffolding the system enhances NADH recycling efficiency, boosting the cascade productivity¹²⁹. Such increase in the cascade productivity is difficult to explain by the apparent catalytic efficiency of each scaffolded enzymes. In particular, the apparent K_M of FDH (rate-limiting enzyme) assembled into SCAB_C towards NAD⁺ (intermediate) is 2.6-fold higher than that of the same enzyme assembled into SCAB_H. Hence, the enzyme apparent kinetics does not support the cascade enhanced performance due to a better proficiency of the rate-limiting enzyme to transform the intermediate when scaffolded in SCAB_H. Hence, we suggest that geometrical and clustering effects must explain the productivity differences found between the two scaffolds. First, we observe that FDH and AlaDH were spatially arranged in a closer distance when scaffolded in SCAB_C than in SCAB_H. However, this difference in inter enzyme distance is too small to explain any kinetic benefit due to close-proximity channeling, according to theoretical studies that state that intermediate concentration gradients are not observed at inter-enzyme distance lower than 1-5 μm under non-restricted diffusion conditions^{3,130}. Another plausible explanation is the formation of an enzyme aggregate driven by an uncontrolled polymerization of SCAB_H modules since the post-translational fusion of one FDH@SCAB1_C dimer and one AlaDH@SCAB2_C hexamer leave 1 SCAB1_C and 5 SCAB2_C unsatisfied, allowing the scaffold aggregation. This hypothesis is explained by large aggregates that are detected both by SDS-PAGE and SEC when FDH@SCAB1_C and AlaDH@SCAB2_C are assembled in solution. The absence of peak in MALDI-TOF analysis also supports the formation of protein aggregates of large molecular weight. In this scenario, when FDH releases NADH, this cofactor will encounter AlaDH in its way heading toward the bulk solution. This type of intermediate channeling is probabilistic and depends more on the size of the aggregate and their concentration in the bulk than on the inter enzyme distance within the scaffold. It has been theoretically reported that for a 2-enzyme scaffold forming aggregates of 260 nm with an optimal separation between aggregates of 6.5 μm, the cascade productivity increases 6 times compared to the non-aggregated and delocalized enzymes. In our case, the productivity enhancement was the half of that theoretically proposed but still hints that SCAB_C aggregates formed due to oligomeric nature of the assembled enzymes are responsible of the enhancement of the overall cascade throughput (Figure 4.16d)¹³¹. The kinetic differences between the two scaffolds in the enzyme cascade could be attributed to the different expected distance between the scaffolded enzymes (Figures 4.12a and 4.12c) and the different supramolecular assembly states detected for the assemblies.

Specifically, the FDH/AlaDH@SCAB_C system showed larger assemblies when compared with FDH/AlaDH@SCAB_H system (Figures 4.12b and 4.12d).

4.3. Conclusions

In conclusion, a straightforward technique for enzyme assembly on a scaffold based on tetratricopeptide consensus repeat (CTPR) proteins; (referred to as SCAB in this work) was developed. We present two modular methodologies that involve selective head-to-tail interactions resulted in the production of a covalent di-sulfide bond and metal-driven assembly techniques as potential tools for the preparation of scaffolded enzyme systems. Through this protein-based scaffolding systems the scaffolded enzymes retained their activity as FDH/AlaDH@SCAB_C and FDH/AlaDH@SCAB_H assemblies, as well as a considerable increase in their catalytic efficiency during biocatalytic processes when assembled as FDH/AlaDH@SCAB_C. This catalytic enhancement can be mainly attributed to the proximity of the scaffolded enzymes in the SCABs with the cysteine-based scaffold (≈ 3.2 nm), behaving more efficiently than those involving the free enzymes. Even though the enzymes were closer to each other in the FDH/AlaDH@SCAB_H (≈ 5.3 nm), the same catalytic enhancement was not achieved probably due to reduced FDH activity after the assembly.

Finally, the previously described scaffolded multi-enzyme system FDH/AlaDH@SCAB_C not only has superior catalytic activity than free enzymes, but it also achieves higher cofactor regeneration rates. This effect is mainly attributed to the closeness of the scaffolded enzymes compared to their free homologous where substrate channeling is not observed. Therefore, this new structured multi-enzyme system enables more effective reuse of the NADH cofactor. This action causes the NADH cofactor to regenerate by effectively passing the byproduct of one enzyme to a nearby enzyme cascade.

The described protein-based scaffolds have the potential to promote the construction of more efficient multi-enzyme systems where scaffolded enzymes are closely allocated. The applications herein demonstrated could be extended to other fields of applied science, e.g. as biosensors for the detection of contaminating analytes¹³².

Chapter 5.
**Engineered repeat proteins as
scaffolds to assemble multi-enzyme
systems for efficient cell-free
biosynthesis**

5.1. Introduction

5.1.1. Cell-free biosynthesis

Cell-free biocatalytic systems are becoming more popular as a substitute for conventional chemical catalysts as enzymes are more sustainable and selective than organic catalysts to manufacture valuable chemicals (Figure 5.1). Applied biocatalysis and more specifically chemical biomanufacturing have benefited from the outstanding advances in molecular and synthetic biology, spurring the creation of novel enzyme cascades^{43,133}. Among other approaches, to increase the throughput of enzyme cascades, biotechnologists have designed protein and DNA-based scaffolds to spatially organize multi-enzyme systems within a few nanometers for efficient cell-free biosynthetic pathways. These scaffolded biocatalytic complexes can be tethered to the surface of artificial materials such as porous agarose beads^{32,134}, cellulose particles¹³⁵, or lipid droplets¹³⁶ to heterogenize them, facilitating their separation and recycling.

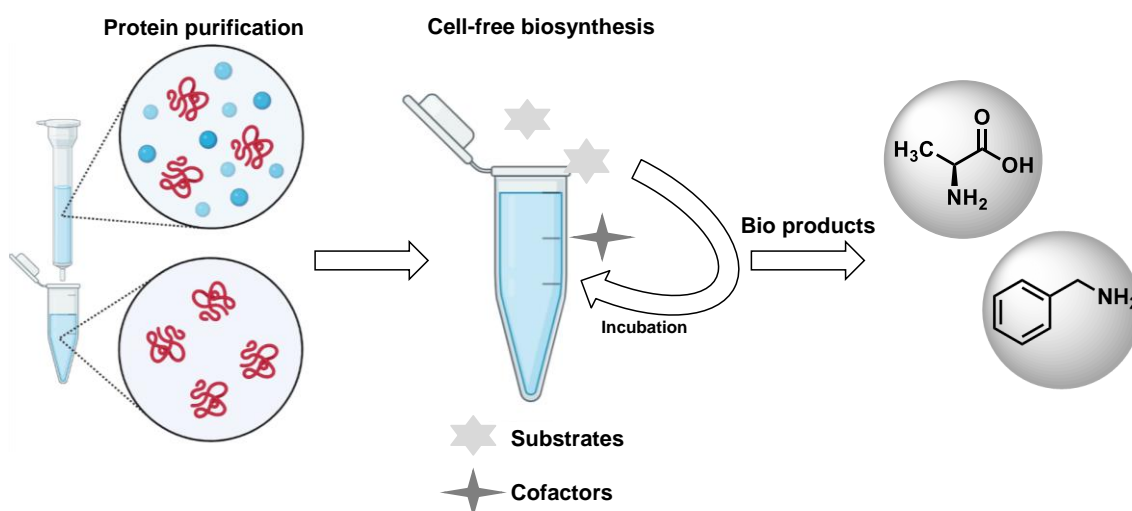


Figure 5.1. Schematic representation of cell-free biosynthesis process. A brief description is shown from the purified proteins to the generation of bio products, as in the case of this thesis the natural amino acid L-Alanine and the organic molecule benzylamine. The intermediate step of the process shows how the catalytic proteins interact with the substrates and cofactors and how the aforementioned bio products already are synthesized.

Among the protein-based scaffolds to organize multi-enzyme systems, cohesin-dockerin interactions derived from natural cellulosomes dominate the landscape of applications^{32,63,64}. Nature offers a limited diversity of cohesin/dockerin pairs, which can be assembled as synthetic cellulosomes formed by fusion proteins containing different cohesin domains (scaffoldin) that reversibly bind several enzymes fused to their cognates dockerin domain through calcium-driven protein-protein interactions with extremely high affinity (K_D in nM-pM range)¹³⁷. Besides cellulosome-based scaffolds, other protein domains like affibodies⁶⁵, leucine zippers⁶⁶, PDZ and SH3 domains^{53,67}, CipB scaffolds⁶⁷, and SpyCatcher/Spy-Tag system⁴³, have been successfully exploited to organize enzymes in space. Both inside and outside of cells, the spatial proximity of these scaffolded enzymes has proven to be an excellent approach to increase the

catalytic efficiency of cascade reactions compared to their non-assembled counterparts where enzymes are diluted in the reaction medium.

5.1.2. Scaffolded enzymes: Spatial arrangement and channeling effect

Although the overall catalytic performance of multi-enzyme systems assembled on biomolecular scaffolds is enhanced, the fundamentals underlying this improvement are still not understood and thus remain the subject of a vibrant debate. The most accepted theory to explain the enhancement in the performance of scaffolded multi-enzyme systems is the improved mass transport of intermediates between the neighboring enzymes. Owing to the physical proximity of enzyme active sites, the transport of intermediates might be facilitated, resulting on a positive impact on the final rate of product formation^{50–52}. This channeling not only benefits the overall kinetics of in vitro cascades but also increases the product titers through reducing the accumulation of labile intermediates in the bulk, avoiding the derailment of intermediates to off-target products and shifting the thermodynamic equilibrium of cascades towards the target product³. Hess and co-workers refuted that nanometric proximity of the enzymes avoids the diffusion of the intermediates to the bulk as the diffusion rate of small molecules is normally much higher than the catalytic efficiency of the enzymes. To explain the greater performance of cell-free multi-enzyme systems when scaffolded within a biomolecular chassis, it has been proposed the existence of a diffusion-limited effect that poses the system out of equilibrium. Such diffusion-limited effect may take place when the intermediates are reversibly absorbed to a scaffolded architecture (*i.e.*, electrostatic interactions). Herein, higher local concentrations of the intermediates in the environment of scaffolded enzymes may occur, enhancing the efficiency of the system when the second enzyme is the rate-limiting one. To a certain extent, these scenarios can be considered a kind of “*intermediate channeling*” that enhances the reaction flux towards the target product (Figure 5.2). The electrostatic guidance emerges as an efficient mechanism for substrate channeling¹³⁸, however experimentally it has only been demonstrated using DNA scaffolds,^{55,139} and by enzyme fusions through positively charged linkers,^{139,140} and up to date has never been addressed using a protein scaffold.

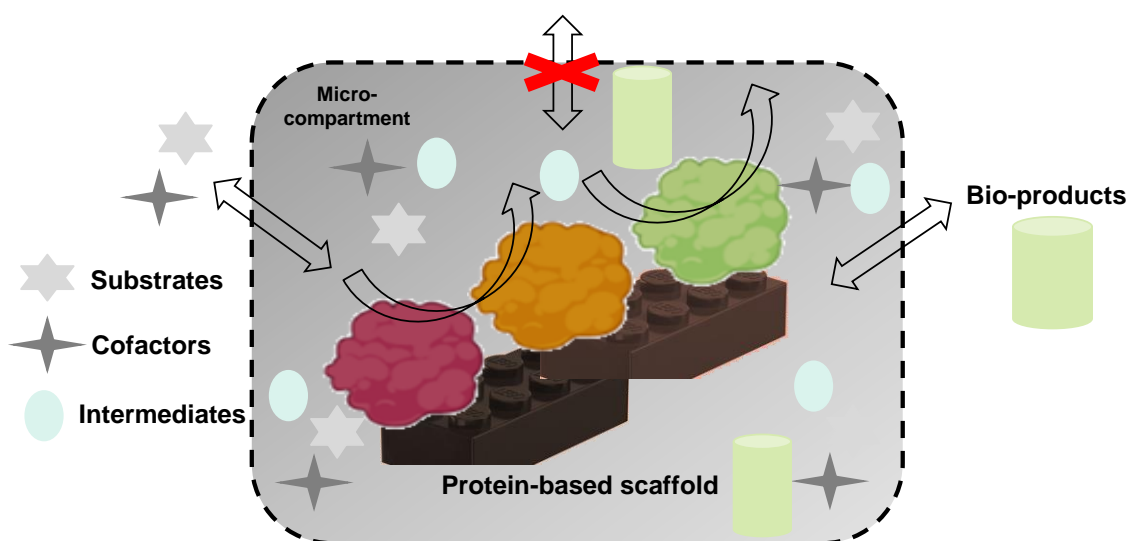


Figure 5.2. The impact of substrate channeling on protein-based scaffolds is depicted schematically. The enzymes in the cascade are arranged on synthetic

scaffolds. The organized enzymes, substrates, and cofactors come into direct contact with one another, resulting in intermediates and, eventually, the desired bioproducts.

5.1.3. TRAP proteins as scaffolding units

Aiming at generating efficient multi-enzyme systems, here we propose to exploit not only the precise protein spatial organization of a multi-enzyme system with nanometric resolution, but also encode interactions with the cascade intermediates to increase their local concentration. Toward this aim, we propose the use of engineered proteins to assemble multi-enzyme systems, encoding by design the nanometric proximity of catalytic domains and the binding sites to reversibly sequester cascade intermediates. Although widely used for the fabrication of genetically programmed biomaterials with multiple applications¹²³, the tetrapeptide repeat proteins (TPR) have not been used as scaffolds to spatially organize multi-enzyme systems. The TPR domain is a 34 amino acid helix-turn-helix repeated motif⁷⁷ that enable protein-protein interactions within the cells. TPR domains are easy to engineer, thereby they can be genetically programmed to mediate protein-protein interactions⁸². Here, we have selected three engineered TPR peptide affinity domains (TRAPs), which bind a set of peptides with high specificity and minimal cross-reactivity inspired by HOP TPR units⁹³. In addition, the physicochemical properties of TPR domains, including surface charge, can be modified by altering non-conserved, surface exposed amino acids¹²⁴. As a result, TRAP-based scaffolds offer greater versatility and flexibility compared to cellulosome-based artificial scaffolds, which have limited natural interaction diversity and are not easily engineerable, as TRAPs can be easily redesigned through interface mutations (Figure 5.3).

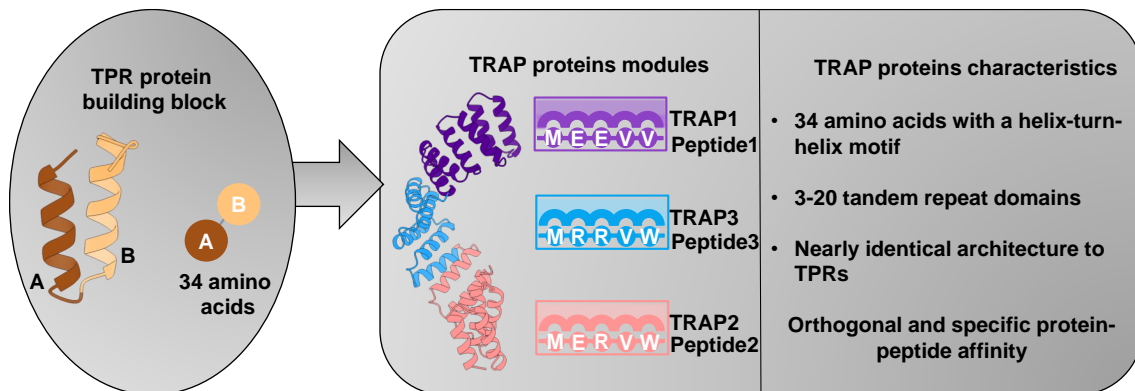


Figure 5.3. Schematic representation of TRAP proteins and its natural counterpart, TPR protein domain. On the left is shown a module of TPR domain consisting of 34 amino acids that encodes a helix-turn-helix motif. On the right, the three engineered TRAP modules (TRAP1, TRAP2, and TRAP3) and their respective engineered affinity peptides (MEEVV, MEVRW, and MRRVW) are shown in purple, pink, and blue, respectively. The most important features of the engineered TRAP proteins are listed, highlighting the most important one: orthogonal and specific recognition of TRAP-peptide pairs.

5.1.4. FDH/ ω TA/AlaDH tri-enzyme cascade on TRAP proteins

Bio-redox reactions coupled to the regeneration of soluble NAD(P)H cofactors are one of the most suitable systems to assess the effects of biomolecular scaffolds on the catalytic efficiency of multi-enzyme systems. These processes work more efficiently when both the main redox reaction and the cofactor recycling step function simultaneously in one-pot as the exogenous cofactor can be supplied at lower concentrations and the thermodynamic limitations due to the underlying reversible nature of the dehydrogenases are minimized. Spatial organization of the participating dehydrogenases is a successful solution to improve the efficiency of *in situ* NADH recycling¹².

In this work, we exploit engineered TRAP domains to assemble a formate dehydrogenase from *Candida boidinii* (FDH)⁹⁵ and an alanine dehydrogenase from *Bacillus stearothermophilus* (AlaDH)⁹⁶. These two enzymes are known to simultaneously perform the asymmetric reduction of α -ketoacids to L-amino acids while recycling NADH using formate as an ancillary electron donor and CO₂ as a by-product. This technology was expanded using a trivalent TRAP to a tri-enzymatic system including the two abovementioned enzymes plus an ω -amino transaminase from *Pseudomonas fluorescens* (ω TA)⁹⁷, for the amination of benzaldehyde with *in situ* recycling of the electron (NADH) and amine (L-Alanine) donors. NADH channeling between the two scaffolded dehydrogenases was demonstrated, resulting in higher cascade productivity and product titer. Finally, the scaffolded system was immobilized on porous solid particles creating a multi-functional and spatially organized heterogeneous biocatalyst that can be recycled for a few cycles. This study provides a solid evidence that the physical proximity of enzymes through assembly on engineered protein scaffolds, along with selective interactions with reaction intermediates, enables intermediate channeling between active sites and results in enhanced catalytic efficiency (Figure 5.4).

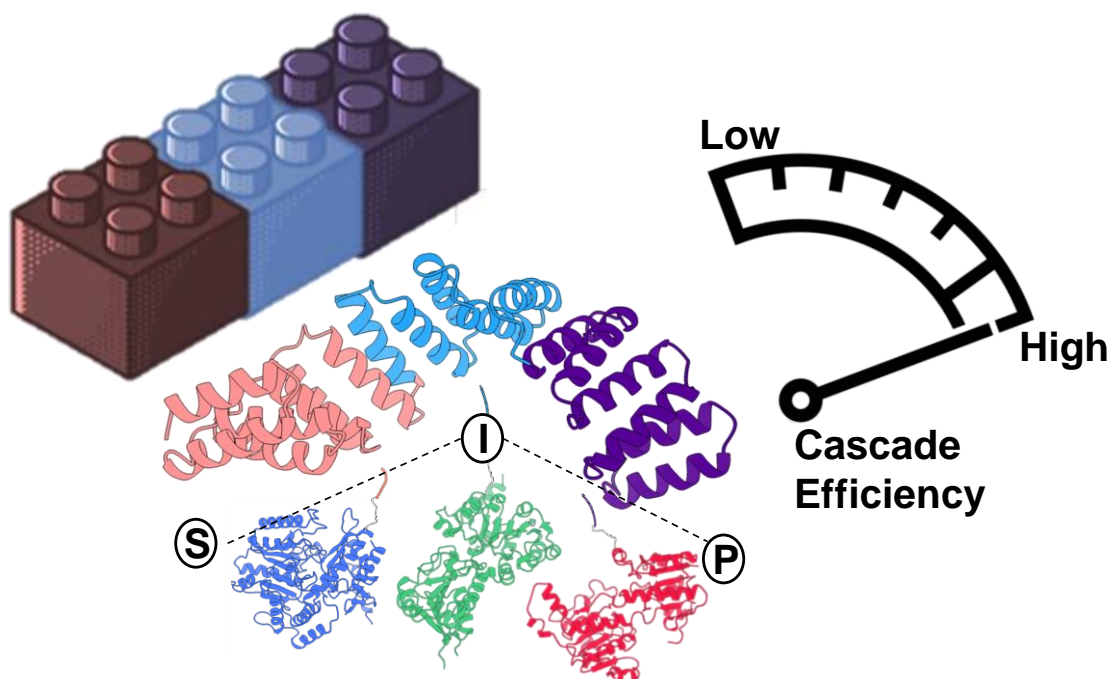


Figure 5.4. Schematic representation of the FDH/ ω TA/AlaDH tri-enzymatic cascade scaffolded on TRAP proteins. This biomolecular recognition-driven scaffolded displayed a highly efficient multi-enzymatic cascade. This phenomenon results in a faster conversion of substrates (S) to products (P), when compared with free enzymes. The intermediates of the cascade reaction are represented as (I). TRAP modules are represented with lego blocks. TRAP1 in purple, TRAP2 in pink, and TRAP3 in blue. The corresponding cognate recognition peptides (MEEVV for TRAP1, MEVRW for TRAP2, and MRRVW for TRAP3) fused to FDH (in red, PDB ID: 5DNA), ω TA (in blue, PDB ID: 5LH9) and AlaDH (in green, PDB ID: 1PJB) were shown.

5.2. Results and discussion

5.2.1. Design of the scaffolding strategy to assemble multi-enzyme systems

Tetratricopeptide Affinity Repeat Proteins (TRAPs) were chosen as scaffolds for enzyme assembly due to their excellent orthogonality in binding different cognate peptides. Initially, we selected two previously designed TRAP1 and TRAP3 modules as they selectively bind the 5-amino acid peptides MEEVV (peptide-1) and MRRVW (peptide-3), respectively (Figure 5.7a)⁹³. TRAP1/TRAP3 pairs are highly orthogonal as the affinity of TRAP1 toward peptide 1 is 250 times tighter than toward peptide 3, while TRAP3 binds to peptide 3 with an affinity 100 times tighter than to peptide 1⁹³. Moreover, the surface exposed residues of TPR domains that are not involved in the peptide binding are negligibly conserved, which means they can be modified without affecting the structure or recognition capabilities. In particular, the surface of TRAP domains exposed patches of positively charged residues that can be exploited to electrostatically bind phosphorylated cofactors, such as NADH and NAD⁺. With this knowledge in mind, we fused peptide-1 (MEEVV), and peptide-3 (MRRVW) to the C-terminus of FDH and AlaDH, respectively (Figure 5.7a). As FDH is a dimer, it displays two units of peptide 1, while the hexameric structure of AlaDH displays six units of peptide 3. Once the FDH and AlaDH enzymes are fused to their respective peptides, they are named FDH1 and AlaDH3, respectively. The TRAP1-3 scaffold was generated by the fusion of the TRAP1 and 3 binding domains to bind nearby the two tagged enzymes FDH1 and AlaDH3 (Table 5.1).

Table 5.1. Amino acid sequences the open reading frame (ORF) for the elements of the assembly: TRAP1-3 and TRAP2-3-1 scaffolds; and tagged-enzymes FDH1, ω TA2, and AlaDH3. TRAP1 binding module in green, TRAP2 binding module in orange, and TRAP3 binding module in purple. Cognate peptides 1 (MEEVV), 2 (MERVW) and 3 (MRRVW) directly fused to FDH, ω TA and AlaDH enzymes in blue.

Assembly elements	Amino acid sequence of the ORFs
TRAP1-3	MGSSHHHHHHSSGLVPRGSHMGSSALKEKELGNDAYKKKDFDTALKHYDKAKELDPTNM YILNQAAVYFEKGDYMKRELCEKAEVGRNREYRLIAIAYARIGNSYFKEEKYKDAIH YKNSLAEHRTPKVLKCCQQAQAEKILKEQGGSLQALKEKELGNDAYKKKDFDTALKHYDK AKELDPTNMYIMNQAAVYFEKGDYMKRELCEKAEVGRNREYRMIAYAYADIGDSYF KEEKYKDAIHFNKSLAEHRTPKVLKCCQQAQAEKILKEQLE
TRAP2-3-1	MGSSHHHHHHSSGLVPRGSHMGSSALKEKELGNDAYKKKDFDTALKHYDKAKELDPTNM YIMNQAAVYFEKGDYMKRELCEKAEVGRNREYRMIAYAYARIGNSYFKEEKYKDAI HFYKNSLAEHRTPKVLKCCQQAQAEKILKEQGGSLQALKEKELGNDAYKKKDFDTALKHY DKAKELDPTNMYIMNQAAVYFEKGDYMKRELCEKAEVGRNREYRMIAYAYADIGDS YFKEEKYKDAIHFNKSLAEHRTPKVLKCCQQAQAEKILKEQGGSLQALKEKELGNDAYKK

Chapter 5. Engineered repeat proteins as scaffolds to assemble multi-enzyme systems for efficient cell-free biosynthesis

	KDFDTALKHYDKAKELDPTNMYILNQAAVYFEKGDYNKCRELCEKAIEVGRENREDYRLI AIAIYARIGNSYFKEEKYKDAIHFNKSLAEHRTPKVLKCCQQAEEKILKEQLE
FDH1	MGSSHHHHHHSSGLVPRGSHMEFKRSMKIVLVLYDAGKHADEEKLYGCTENKLGIAN WLKDQGHELITTSDEKEGETSELDKHIPDADIIITPPFHPAYITKERLDAKAKNLKLVVVAGVG SDHIDLDIYINQTKKISVLEVTGSNVVSVAEHVVMTMLVLRNFVPAHEQIINHWEVAAI AKDAYDIEGKTIATIGAGRIGYRVLERLLPFNPKELLYDYQALPKEAEEKVGGARRVENIEE LVAQADIVTVNAPLHAGTKLINKELLSKFKKGAWLVNTARGAICVAEDVAAALESGQLR GYGGDVVFPQPAPKDPWRDMRNKYGAGNAMTPHYSGTTLDAQTRYAEGTKNILESF FTGKFDYRPQDIILLNGEYVTKAYGKHDKKKKLGSGSGSGSDDTSR MEEVV
ωTA2	MGSSHHHHHHSSGENLYFQGHMLRSNSNNKAWLKEHNTVHMMHPMQDPKALHEQRP LIIQSGKGVHITDVGRRFIDCQGGWLWCVNAGYGRREIIDAVTRQMEELAYYSLFPGSTN APAIALSQKLTEVAAEEGMVKASFGLGGSDAVETALKIARQYWKLEGQDPKV/KFVSLYN GYHGLNFGGMSACGGNAWKSSYEPLMPGFFQVESPPLYRNPFTNDPEELAEICAQILER QIEMQAPGTVAALIAEPIQGAGGVVPPASYWPRLRQICDKYDILLIADDEVITGLGRSGSLF GSRGWGVKPDIMCLAKGISSGYVPLSATLVNSRVARAWERDAGFTSVYMHGYTYSGHP VSCAAALAAIDIVLQENLAENARVVGDFLEKLLILKDKHRAIGDVRGKGLMLLAVELVKER ATKEPFGPADAYPLAISEACVNNGVMIRTIVNKLIISPPLTFTTEHVDEVIEVLDRAFVANP WKLGSAGSGSDDTSR MERVV
AlaDH3	MGSSHHHHHHSSGLVPRGSHMEFKRSIIGVPKEIKNNENRVALTPGGVSQLISNGHRVL VETGAGLGSGFENEAYESAGAEIADPKQVWDAEMVMKVKEPLPEEYVYFRKGLVLFY LHLAAPELAQALKDKGVTAIAYETVSEGRTLPLLLTPMSEVAGRMAAQIGAQFLEKPKGG KGILLAGVPGVSRGKVTIIGGGVGTNAAKMAVGLGADVITIDLNADRLRQLDDIFGHQIKT LISNPVNIADAVAEADLLICAVLIPGAKAPTLVTEEMVKQMKPGSVIVDVAIDQGGIVETVD HITTHDQPTYEKHGVVHYAVANMPGAVPRTSTIALTNVTVPYALQIANKGAVKALADNTAL RAGLNTANGHVTYEAVARDLGYEYVPAEKALQDESSVAGAKLGSAGSGSDDTSRMDDT RMRRVV

These protein constructs were expressed, purified, and characterized. First, their molecular masses were analyzed by electrophoresis (SDS-PAGE) and mass spectrometry. As expected, based on their sequence, TRAP1-3 scaffold presented a mass of 32.621 kDa, FDH1 a mass of 45.137 kDa, and AlaDH3 a mass of 44.323 kDa (Figure 5.5a and 5.5b). Additionally, circular dichroism analysis revealed that the secondary structure of the tagged enzymes was negligibly affected by the peptide fusion and showed that the individual proteins were stable and displayed cooperative thermal denaturation transitions (Figure 5.5c).

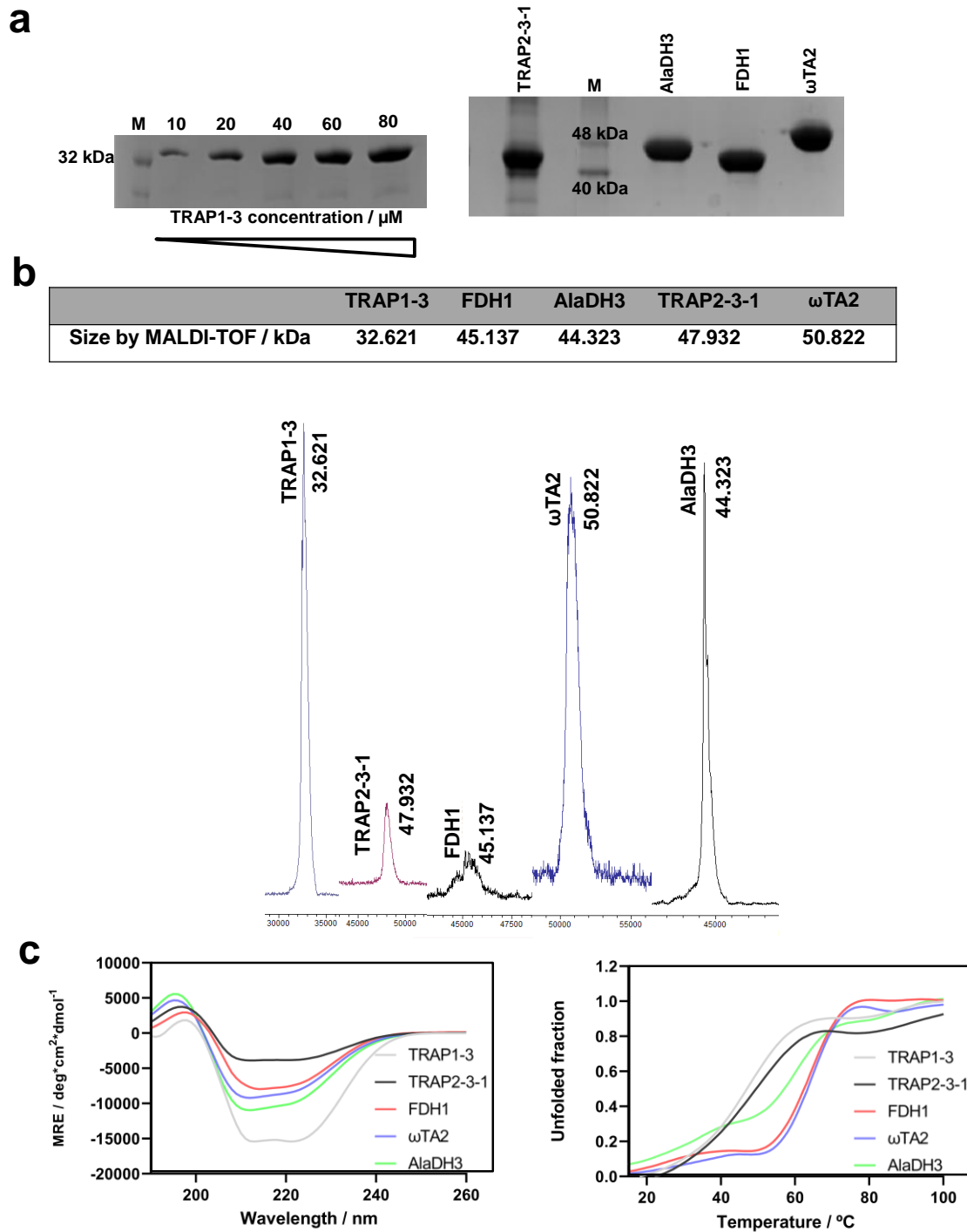


Figure 5.5. Characterization of TRAP1-3 and TRAP2-3-1 scaffolds and tagged enzymes (FDH1, ωTA2 and AlaDH3). a) SDS-PAGE gel electrophoresis of purified TRAP scaffolds and tagged enzymes. Left panel: TRAP1-3 scaffold in different μM concentrations. Right panel: TRAP2-3-1 scaffold and tagged enzymes, AlaDH3, FDH1 and ωTA2. b) MALDI-TOF spectra and calculated masses from the MALDI spectra of the purified proteins. TRAP1-3 scaffold: 32.621 kDa, TRAP2-3-1 scaffold: 47.932 kDa, FDH1: 45.137 kDa, ωTA2: 50.822 kDa and AlaDH3: 44.323 kDa tagged enzymes. Masses calculated by the amino acid composition for TRAP1-3 scaffold: 32.559 kDa, TRAP2-3-1 scaffold: 47.803 kDa, FDH1: 45.276 kDa, ωTA2: 50.433 kDa and AlaDH3:

44.471 kDa tagged enzymes. c) Left panel, circular dichroism spectra of TRAP1-3 and TRAP2-3-1 scaffolds, FDH1, ω TA2 and AlaDH3 tagged enzymes. Right panel, thermal denaturation curves monitored by the decrease of the CD signal at 222 nm of TRAP1-3 and TRAP2-3-1 scaffolds, FDH1, ω TA2 and AlaDH3 tagged enzymes. TRAP1-3 in grey, TRAP2-3-1 in black, FDH1 in red, ω TA2 in dark blue, and AlaDH3 in green.

The oligomeric state of the enzymes was evaluated by SEC. The estimated MW by SEC for AlaDH3 and FDH1 were 286 kDa and 95 kDa, respectively, in agreement with the expected hexameric state of AlaDH3 (266 kDa), and dimeric state of FDH1 (90 kDa) (Figure 5.6 and Table 5.2).

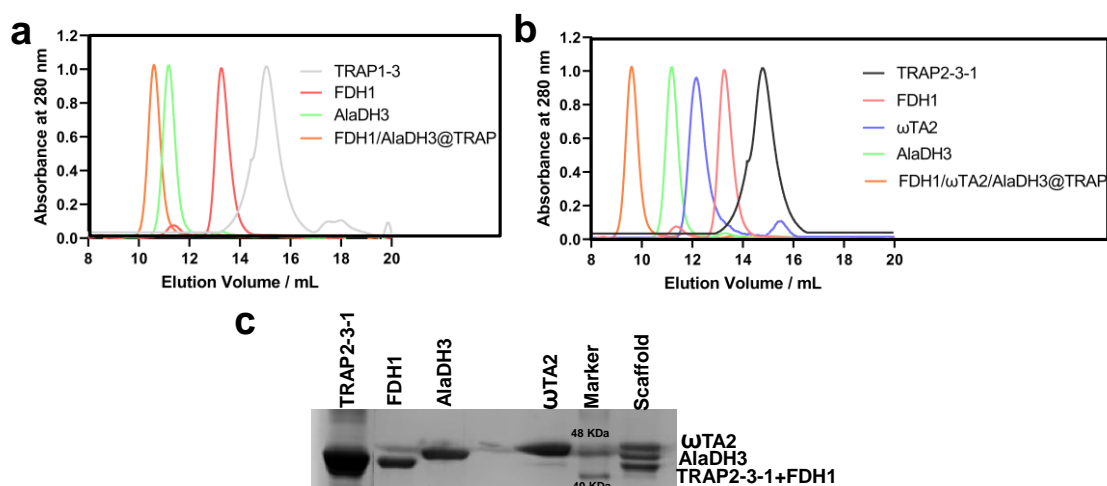


Figure 5.6. Characterization of the two and three-enzyme assemblies based on biomolecular recognition by SEC using Superdex 200 (10/300) column. a) SEC chromatogram monitored at 280 nm for TRAP1-3 scaffold, FDH1 and AlaDH3 tagged enzymes and the scaffolded enzyme system, FDH1/AlaDH3@TRAP. b) SEC chromatogram monitored at 280 nm for TRAP2-3-1 scaffold, FDH1, ω TA2 and AlaDH3 tagged enzymes and the scaffolded enzyme system, FDH1/ ω TA2/AlaDH3@TRAP. c) SDS-PAGE gel of the purified elements after SEC analysis. From left to right: the free elements TRAP2-3-1, FDH1, AlaDH3 and ω TA2, protein marker and scaffold system FDH1/ ω TA2/AlaDH3@TRAP. The FDH1/ ω TA2/AlaDH3@TRAP composite elements were separated by size from largest to smallest: ω TA2, AlaDH3 and TRAP2-3-1+FDH1. The latter cannot be separated because their size in the gel is too similar.

Table 5.2. Quantitative size analysis of the single component and scaffolded enzyme systems by SEC using a Superdex 200 (10/300) column. The elution volumes for each TRAP protein, tagged enzyme, and scaffolded enzyme system are shown. The MW of the different systems was estimated from the elution volume by using a calibration curve ($y = -0.213x + 7.816$) generated with gel filtration calibration kit composed of a mixture of well-defined proteins standards. The MW estimated experimentally is compared with that obtained from the amino acid sequence.

Sample	Elution volumen / mL	MW / kDa (by SEC)	MW / kDa (amino acid composition)
TRAP1-3	15.31	35	32
TRAP2-3-1	14.70	48	47
FDH1 (dimer)	11.21	95	90
ω TA2 (tetramer)	11.99	183	200
AlaDH3 (hexamer)	13.32	268	266
FDH1AlaDH3@TRAP	10.85	370	388 (1:1:1)
FDH1/ ω TA2/AlaDH3@TRAP	9.55	605	603

Therefore, the tagging of cognate peptides affected neither the secondary nor the quaternary structure of the enzymes. The specific enzyme activity of the FDH and AlaDH was evaluated and compared to the enzymes fused to the recognition peptides (FDH1 and AlaDH3) showing that the addition of the peptides negligibly affects their specific activity (Table 5.3).

Table 5.3. Specific enzymatic activity of the enzymes upon peptide fusion. Specific enzymatic activities / U·mg⁻¹ of the FDH, ω TA and AlaDH enzymes compared to their corresponding tagged-enzymes, FDH1, ω TA2 and AlaDH3.

Enzymes	Specific enzymatic activity / U·mg ⁻¹
FDH	0.82 ± 0.07
FDH1	0.74 ± 0.11
ω TA	5.58 ± 0.10
ω TA2	2.38 ± 0.05
AlaDH	22.1 ± 0.01
AlaDH3	20.8 ± 0.33

Once the scaffoldin unit and the tagged enzymes were structurally and functionally characterized, we performed their stepwise assembly in solution. First, TRAP1-3 was mixed with AlaDH3 for 1 hour at 4 °C, and then such one-enzyme assembly was incubated with FDH1 under the same conditions to form the target two-enzyme assembly (Figure 5.7b). As the assembly was performed sequentially, AlaDH3 hexamers were used as the nucleating unit of the assembly and incubated with TRAP1-3 monomers at a stoichiometry of 1:6 to load all the AlaDH3 subunits with one TRAP unit. Then, three FDH1 dimers were incubated per hexamer of AlaDH3, resulting in an assembly with a theoretical stoichiometry of 6:3:1 (TRAP1-3:FDH1:AlaDH3), which means a 1:1:1 molar ratio, when accounting for the protein monomers. In this compact and well-defined assembly, all the tagged peptides are expected to be bound to their corresponding TRAP module (Figure 5.7b). To evaluate the assembly process TRAP1-3 scaffold was labelled with AF647 dye. The apparent Rh of the labeled TRAP1-3 domain, and the changes in the apparent hydrodynamic radius (Rh) upon enzyme assembly was assessed by microfluidic diffusional sizing¹⁰⁴. The TRAP1-3:FDH1:AlaDH3 complexes showed an increase of their Rh compared to TRAP1-3. The apparent Rh of the assembled system was larger than the two non-scaffolded enzymes and the tagged enzymes individually

assembled into TRAP1-3 (Figure 5.7c). These results demonstrate the efficient assembly of the two tagged enzymes within the engineered scaffold.

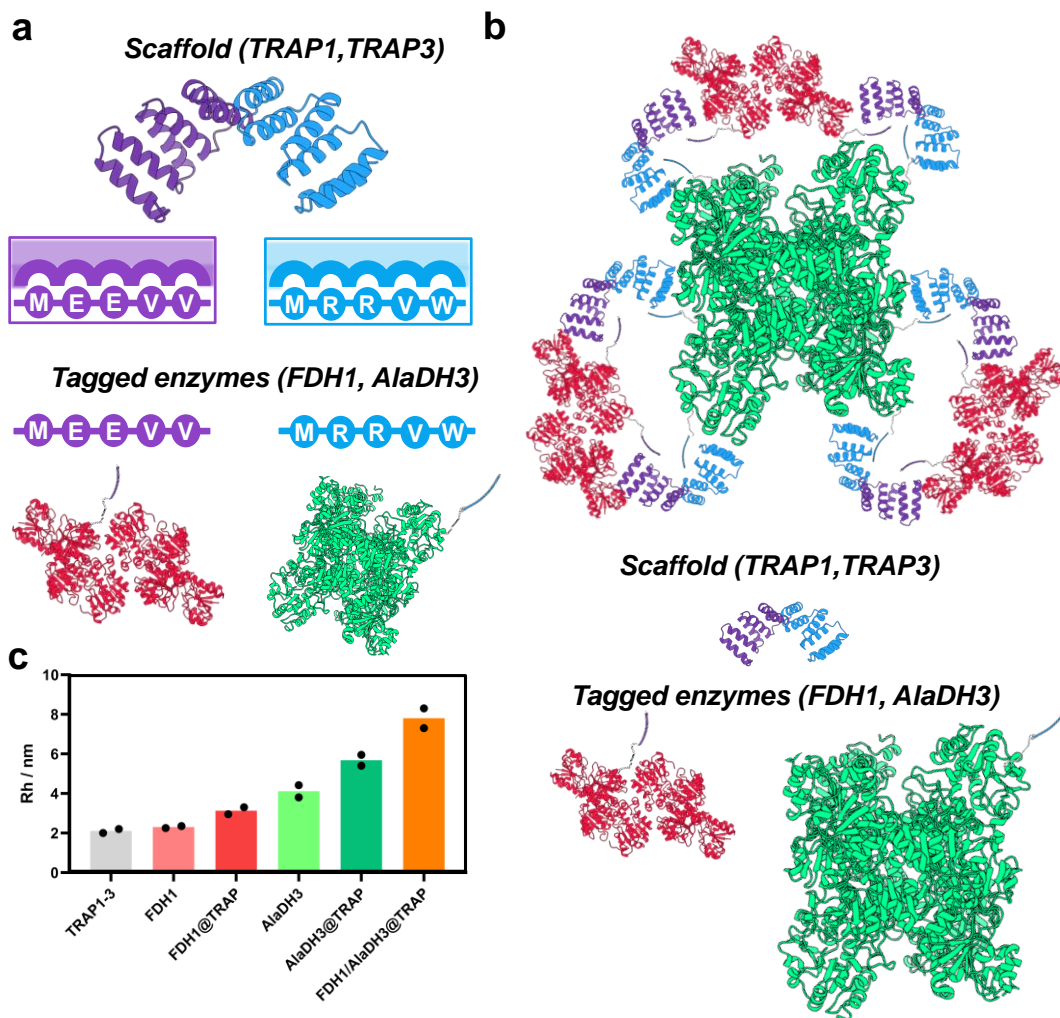


Figure 5.7. Scaffolding strategy to assemble multi-enzymatic systems. a) Schematic representation of the TRAP-based scaffold composed by TRAP1 (in purple) and TRAP3 (in blue) and their corresponding cognate recognition peptides (MEEVV for TRAP1 and MRRVW for TRAP3) fused to FDH dimer (in red, PDB ID: 5DNA [<https://doi.org/10.2210/pdb5DNA/pdb>]) and AlaDH hexamer (in green, PDB ID: 1PJB [<https://doi.org/10.2210/pdb1PJB/pdb>]). b) Stepwise assembly concept, in which first AlaDH3 hexamer nucleates 6 TRAP1-3 units and the complex is then loaded with three FDH1 dimers, resulting in a 6:3:1 TRAP1-3:FDH1:AlaDH3 stoichiometry in the final assembly (*i.e.* a 1:1:1 stoichiometry of monomers). c) The apparent Hydrodynamic radius (Rh) of the different systems measured by Fluidity One W. The apparent Rh of free scaffold (TRAP1-3), tagged enzymes (FDH1 and AlaDH3), incomplete assemblies with only one enzyme bound (FDH1 @TRAP and AlaDH3@TRAP), and complete assembly with the two enzymes bound (FDH1/AlaDH3@TRAP). The data are presented as the mean of two replicate experiments ($n = 2$).

Once the assembly conditions were optimized, we analyzed the assemblies by SEC to determine the components of the assembly and its approximate molecular weight, and

the approximate overall molecular weight of the protein assembly. The FDH1/AlaDH3@TRAP assembly showed a unique chromatographic peak indicating the absence of free elements and incomplete assemblies, and the efficient engagement of TRAP scaffold in the interactions with the different subunits of the oligomeric enzymes. The resulting assembly corresponds to a molecular mass of approx. 370 kDa, consistent with the expected assembly and the 6:3:1 stoichiometry utilized in the incubation (Figure 5.7b, Table 5.2). The SEC results confirmed the assembly of the three elements and the changes observed for the apparent R_h upon the stepwise enzymatic assembly. These results suggest that TRAP1-3 scaffold can simultaneously assemble both FDH1 and AlaDH3 enzymes in a 3:1 stoichiometry, where the recycling dehydrogenase is excess regarding the main one.

5.2.2. Catalytic activity of the multi-enzymatic system

The kinetics of free and scaffolded enzymes were evaluated based on UV-Vis assays using different substrates depending on the enzyme activity to be measured. While FDH1 oxidizes formate to CO_2 concomitantly reducing NAD^+ to NADH, AlaDH3 catalyzes the reduction of that NADH to perform the reductive amination of pyruvate in presence of ammonium chloride to yield L-Alanine as the target product. Both enzyme activities were determined by monitoring either the consumption or the formation of NADH at 340 nm. When FDH1 and AlaDH3 are individually bound to TRAP1-3, they maintained more than 80 % activity compared to the activity of their free counterparts (Figure 5.8). Furthermore, when both enzymes were assembled into TRAP1-3 scaffold, their specific activities were also unaffected indicating that the scaffolding process keeps the functionality of the assembled enzymes.

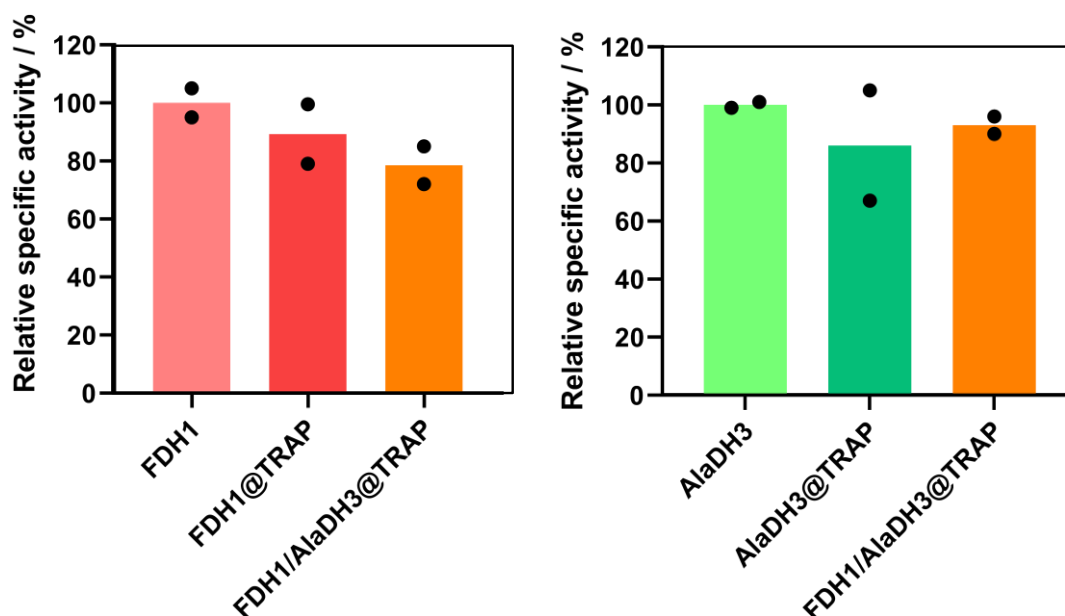


Figure 5.8. Relative activity of scaffolded enzyme systems on TRAP assemblies. Relative activity of free enzymes (FDH1 and AlaDH3) and scaffolded enzyme systems (FDH1@TRAP, AlaDH3@TRAP and FDH1/AlaDH3@TRAP). 100 % of AlaDH and FDH activities correspond to $22.1 \text{ U}\cdot\text{mg}^{-1}$ and $0.83 \text{ U}\cdot\text{mg}^{-1}$, respectively. The data are presented as the mean of two replicate experiments ($n = 2$).

Chapter 5. Engineered repeat proteins as scaffolds to assemble multi-enzyme systems for efficient cell-free biosynthesis

To further investigate the catalytic performance of the scaffolded enzymes, we measured the apparent Michaelis-Menten kinetic parameters of the two enzymes either individually scaffolded (FDH1@TRAP and AlaDH3@TRAP) or assembled together (FDH1/AlaDH3@TRAP). As both FDH1 and AlaDH3 are multi-substrate enzymes, we determined the Michaelis-Menten kinetics by varying the concentration of one substrate and fixing saturating concentrations of the others. From the Michaelis-Menten plots (Figure 5.9), we determined K_M , V_{max} , k_{cat} , and k_{cat}/K_M values for all biocatalysts toward all the substrates (Table 5.4).

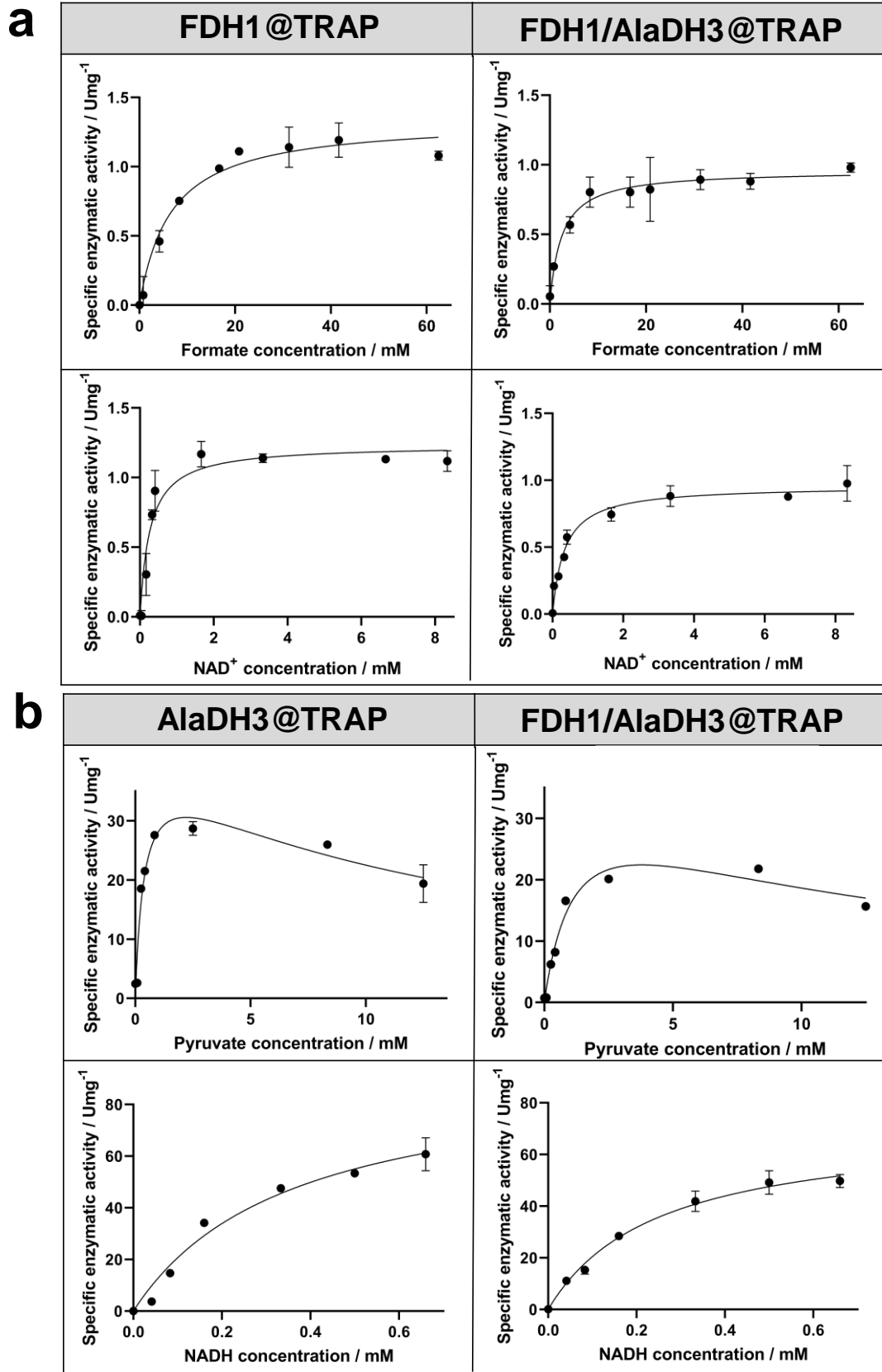


Figure 5.9. Michaelis-Menten curves of scaffolded enzyme systems. a) Michaelis-Menten curves of the scaffolded enzyme systems (FDH1@TRAP and FDH1/AlaDH3@TRAP) using different substrates: variable formate concentration at 1

mM NAD⁺ in 25 mM sodium phosphate buffer pH 7; and variable NAD⁺ concentration at 100 mM formate in 25 mM sodium phosphate buffer pH 7. b) Michaelis-Menten curves of the scaffolded enzyme systems (AlaDH3@TRAP and FDH1/AlaDH3@TRAP) using different substrates: variable pyruvate concentration at 500 mM ammonium chloride, 0.5 mM NADH in 25 mM potassium phosphate buffer pH 8; and variable NADH concentration at 75 mM pyruvate, 500 mM ammonium chloride in 25 mM potassium phosphate buffer pH 8. Pyruvate kinetics were adjusted to a substrate inhibition model. The data are presented as the mean of three replicate experiments (n = 3), and error bars represent standard deviations.

Table 5.4. Michaelis-Menten steady-state kinetic parameters. Enzyme kinetic parameters for FDH as FDH1@TRAP and FDH1/AlaDH3@TRAP against the different substrates (formate and NAD⁺). Enzyme kinetic parameters for AlaDH as AlaDH3@TRAP and FDH1/AlaDH3@TRAP against the different substrates (pyruvate and NADH). k_{cat} were calculated using the MW of the monomers of FDH1 and AlaDH3.

Enzymatic system	Formate			
	K_M / mM	V_{max} / $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$	k_{cat} / s^{-1}	k_{cat}/K_M / $\text{M}^{-1}\cdot\text{s}^{-1}$
FDH1@TRAP	2.42 ± 1.07	0.95 ± 0.20	0.7 ± 0.2	2.9 × 10 ²
FDH1/AlaDH3 @TRAP	6.61 ± 1.71	1.34 ± 0.55	1.0 ± 0.4	1.5 × 10 ²
Enzymatic system	NAD ⁺			
	K_M / mM	V_{max} / $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$	k_{cat} / s^{-1}	k_{cat}/K_M / $\text{M}^{-1}\cdot\text{s}^{-1}$
FDH1@TRAP	0.27 ± 0.10	1.23 ± 0.34	0.9 ± 0.3	3.5 × 10 ³
FDH1/AlaDH3 @TRAP	0.34 ± 0.15	0.96 ± 0.59	0.7 ± 0.5	2.1 × 10 ³
Enzymatic system	Pyruvate			
	K_M / mM	V_{max} / $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$	k_{cat} / s^{-1}	k_{cat}/K_M / $\text{M}^{-1}\cdot\text{s}^{-1}$
AlaDH3@TRAP	0.39 ± 0.05	41.31 ± 2.52	30.6 ± 1.9	7.9 × 10 ⁴
FDH1/AlaDH3 @TRAP	1.25 ± 0.55	37.20 ± 3.34	27.6 ± 2.5	2.2 × 10 ⁴
Enzymatic system	NADH			
	K_M / mM	V_{max} / $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$	k_{cat} / s^{-1}	k_{cat}/K_M / $\text{M}^{-1}\cdot\text{s}^{-1}$
AlaDH3@TRAP	0.36 ± 0.07	94.27 ± 6.39	69.9 ± 4.7	1.9 × 10 ⁵

FDH1/AlaDH3	0.25 ± 0.14	71.15 ± 8.01	52.7 ± 5.9	2.1 × 10 ⁵
@TRAP				

The Michaelis-Menten kinetic parameters clearly show that the rate-limiting step enzyme of this bioredox cascade is the FDH1 as its k_{cat} is two orders of magnitude lower than the k_{cat} of the AlaDH3. The kinetic data evidence that k_{cat} is barely affected by scaffolding the two enzymes together, while K_M values tend to increase when the two enzymes are anchored together within the TRAP1-3 scaffold. The increase in K_M leads to a decrease in the catalytic efficiency (k_{cat}/K_M). Remarkably, the increase in K_M was more significant for the smaller substrates such as formate (MW 46) and pyruvate (MW 87) than for the bulkier redox cofactors NAD⁺ (MW 663) and NADH (MW 664). This increase in K_M indicates that formate and pyruvate have lower affinity for the catalytic centers of the FDH1 and AlaDH3, respectively, when both enzymes are assembled into the TRAP1-3 scaffold. Small substrates such as formate and pyruvate are known to travel toward the enzyme active sites through narrow tunnels embedded in the protein structure^{141,142}. In contrast, NAD cofactors are normally bound to dehydrogenases in a more open conformation, inducing a conformational change that facilitates the productive binding of small substrates (aldehydes, ketones, or alcohols). The different access pathways of the two substrates involved in the catalytic mechanism of these two alcohol dehydrogenases may explain why K_M for the small substrates was affected to a higher extent than for NAD cofactors when the enzymes are scaffolded. In the assembled state, the narrower tunnels to allocate either the pyruvate or the formate in the enzyme active sites may be partially compromised by steric impediments that emerged upon the enzyme assembly. Despite the higher K_M , the recovered activity upon assembly is still high enough to test this system in a model bioredox cascade. This cascade will transform pyruvate into L-Alanine using ammonium chloride as co-substrate for the reductive amination catalyzed by the AlaDH3 and use formate as the ancillary electron donor for the *in-situ* recycling of NADH, which is orthogonally catalyzed by FDH1.

5.2.3. Biosynthesis of L-Alanine catalyzed by TRAP-scaffolded enzyme system

To assess the benefits of the scaffold in the kinetic performance of the multi-enzyme system, we incubated the scaffolded and free biocatalytic systems with pyruvate, 1.3 equivalents of sodium formate, 6.6 equivalents of ammonium chloride and substoichiometric amounts of NADH (150-fold less than pyruvate) (Figure 5.10a). Figure 5.10b shows that the scaffolded system at 1:1:1 molar monomer ratio of TRAP1-3:FDH1:AlaDH3 is five times faster than the free one at the same enzyme stoichiometry, reaching a specific productivity of 5.21 g × g_{enzyme}⁻¹ × h⁻¹ and a titer of 49 mM of L-Alanine after 24 hours, which means a chromatographic yield (CY) of 65 %. In contrast, free FDH1 and AlaDH3, resulted in a CY of only 19 % under the same conditions and incubation time. Considering that the rate-limiting enzyme of this cascade is FDH1 (Table 5.4), we suggest that scaffolding the system improves the NADH recycling efficiency and thus increases the apparent FDH1 activity. Hence, we hypothesize that the higher specific productivity and product yield achieved with the scaffolded system are the results of a more efficient transport of NAD species between the two proximal enzymes.

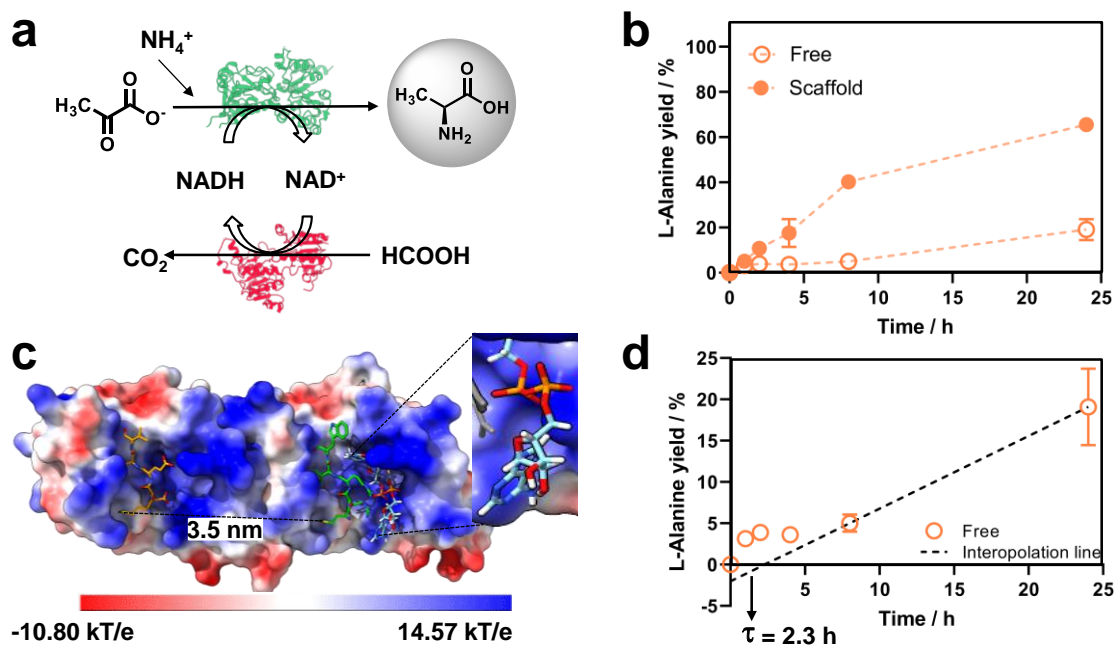


Figure 5.10. L-Alanine synthesis activity of the scaffolded enzymes in batch-mode. a) Scheme of the catalytic cycle that was tested by the HPLC. b) L-Alanine yield at 1:1 molar ratio of FDH1:AlaDH3 monomers for the free enzyme system (FDH1/AlaDH3, empty circles) and the scaffolded enzyme system (FDH1/AlaDH3@TRAP, full circles). c) Structure of the TRAP1-3 scaffold showing the surface electrostatic potential. Peptides 1 and 3 are shown in orange and green sticks, respectively and the distance between the two recognition sites is shown. NADH co-factor docked on the TRAP scaffold is shown in light blue sticks representation. d) Interpolation line for FDH1/AlaDH3 free system with the transient time highlighted with an arrow. Reaction mixture: 100 mM formate, 75 mM pyruvate, 500 mM ammonium chloride and 0.5 mM NADH. The data are presented as the mean of two replicate experiments ($n = 2$), and error bars represent standard deviations.

To further support this hypothesis, we incubated the system with a range of NADH concentrations (0.01-1 mM). We found the maximum performance of the scaffolded system at 0.5 mM NADH (Figure 5.11).

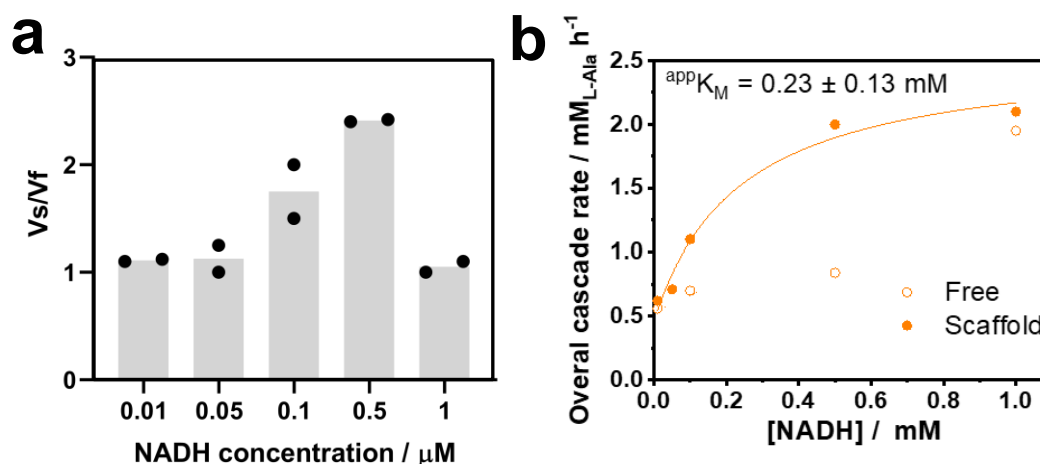


Figure 5.11. a) Ratio of the reaction rate of free (V_f) (FDH1/AlaDH3) and scaffolded (V_s) (FDH1/AlaDH3@TRAP) enzyme systems as a function of NADH concentration. Range of NADH concentrations evaluated: 0.01 mM, 0.05 mM, 0.1 mM, 0.5 mM, and 1 mM at an end point of 4 h. **b)** Saturation plot of the overall enzyme cascade at different NADH bulk concentrations. The apparent K_M value indicated within the plot refers to the apparent K_M of the cascade towards NADH upon fitting the rate datapoints to the Michaelis-Menten equation of the FDH1/AlaDH3@TRAP enzyme system. The data are presented as the mean of two replicate experiments ($n = 2$).

The positive effect provided by the spatial proximity disappears at higher and lower NADH concentrations as the FDH1 may be already saturated at 1 mM NADH in the bulk (three times higher than its K_M) or, contrarily, it is too limited by NADH bulk concentrations < 0.05 mM (six times lower than its K_M). The greatest titer of L-Alanine was also achieved at 0.5 mM, which supports this NADH concentration as the optimal one to maximize the positive effects of the scaffolding on the throughput of the biocascade (Table 5.5).

Table 5.5. L-Alanine conversion as a function of the NADH concentration for FDH1/AlaDH3 and FDH1/AlaDH3@TRAP. The range of NADH cofactor analyzed was from 0.01 mM to saturated conditions; 1 mM.

NADH concentrations / mM	L-Alanine concentration / mM	
	FDH1/AlaDH3	FDH1/AlaDH3@TRAP
0.01	13.5	15.0
0.05	14.8	17.1
0.1	16.8	26.0
0.5	20.29	49.0
1	46.98	50.0

Furthermore, the reaction courses shown in Figure 5.10b indicate the existence of a lag in L-Alanine production using the free enzymes. The transient time of the free system was calculated to be 2.3 hours by linearly fitting the reaction course with the steepest slope from time 8 to 24 hours (Figure 5.10b). In contrast, the scaffolded system exhibited no transient time, indicating that the free system takes longer to reach its maximum

steady-state production rate compared to the scaffolded one, which reaches maximum throughput at the beginning of the reaction.

To illustrate the spatial proximity of the scaffolded enzymes, we generated a molecular model based on the crystallographic structures of FDH1, AlaDH3, and TRAP1-3 (Figure 5.10c). In this model the distance between assembled FDH1 and AlaDH3 is 3.5 nm. In contrast, at the assay protein concentration (0.18 μM), the free enzymes would be separated by a distance of approximately 209 nm according to equation 1 reported by Ellis et al.³.

$$(1) d_{e-e} = \frac{1.18}{C^{1/3}}$$

where d_{e-e} is the distance between the free enzymes and C is the concentration of the free enzymes. Considering the inter-enzyme distances (d_{e-e}) of both the scaffolded and the free systems, and the hydrodynamic radius of the dimmer FDH1 as rate-limiting enzyme ($r = 2.3$ nm), the fraction of NADH that could be directly channeled from FDH1 to AlaDH3 (f_{direct}) is 67 % and 1 %, respectively, according to the equation 2 also presented by Ellis et al.³ Thus, the proximity of the two enzymes promoted by the scaffold increases the chances of NAD species to find their next enzyme.

$$(2) f_{direct} = \frac{r}{d_{e-e}}$$

Although the scaffolded system can theoretically channel 67 % of NADH formed by the FDH1 to AlaDH3 due to physical proximity (see the previous section), we suggest that the cascade is not directly benefited through channeling by proximity according to Wheeldon et al.¹³⁰ Pure diffusion of NADH from one enzyme to the other is hardly plausible since the diffusion coefficient of this redox cofactor in aqueous media ($4 \times 10^6 \text{ cm}^2 \text{ s}^{-1}$)¹⁴³ is much larger than the catalytic efficiency of the rate-limiting enzyme ($3 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$); FDH1 bound to TRAP1-3. Instead, we hypothesize that channeling assisted by electrostatic interactions may facilitate the transport of NADH between the two scaffolded dehydrogenases. Surface charge calculations revealed that the surface of TRAP1-3 contains positively charged patches that may interact electrostatically with the negatively charged phosphate groups of the NAD species. Docking studies demonstrate that both reduced and oxidized cofactors can bind to a positively charged cavity within TRAP1-3 (Figure 5.10c and Figure 5.12).

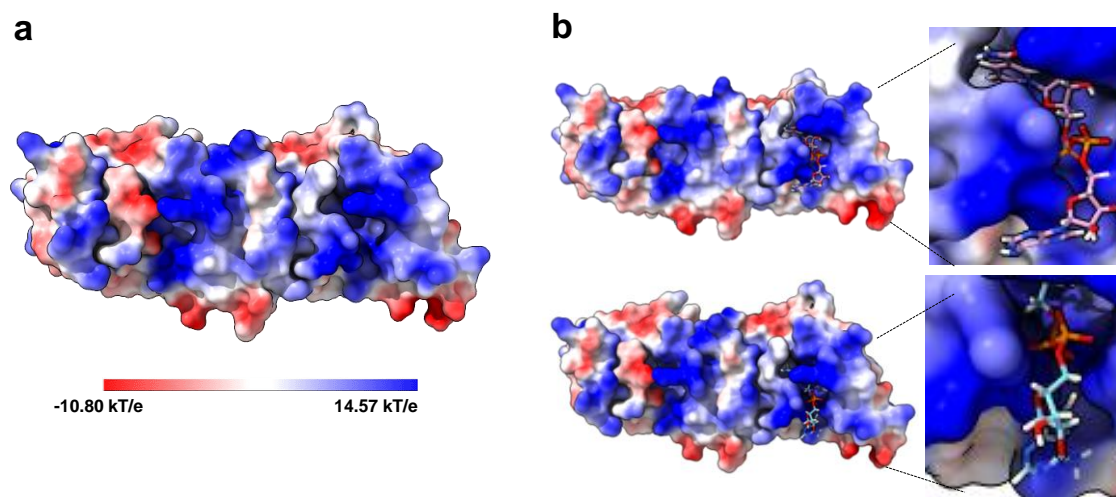


Figure 5.12. Surface charge calculations of TRAP1-3 and molecular docking of TRAP1-3 scaffold and NAD⁺/NADH cofactors. a) Electrostatic surface potential of TRAP1-3 protein scaffold colored according to calculated electrostatic potential of accessible surface area from -10.80 kT/e (red) to 14.57 kT/e (blue). b) Molecular docking studies on the scaffold-cofactor interaction. Molecular models for TRAP1-3 scaffold shown in surface charge representation, and NAD⁺ (top) and NADH (bottom) cofactors shown in sticks representation in light pink and light blue, respectively. Both cofactors interact with the same specific positively charged pocket located in TRAP3 protein module.

This interaction is further supported by an anisotropy-based assay that exploits the intrinsic fluorescence of NADH and results in a TRAP1-3 NADH interaction with a binding constant of K_D 34.9 μ M (Figure 5.13). Similar electrostatic interactions between the scaffolds and the intermediates have been reported for DNA scaffolds and bi-enzyme systems fused through positively charged linkers^{139,140}. Likewise, we suggest that electrostatic interactions between the intermediate (NAD) and the TRAP1-3 scaffolding unit support the channeling we observe.

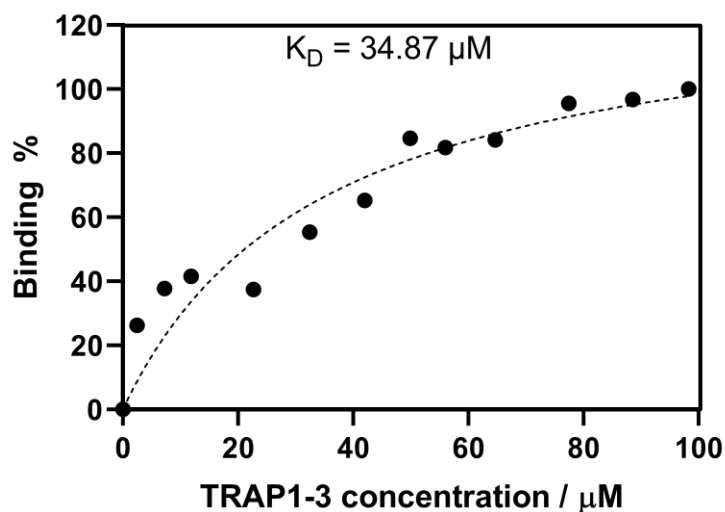


Figure 5.13. Fluorescence anisotropy-based binding assay of the NADH cofactor to the TRAP1-3 scaffold. The binding curve shows the % binding of NADH to different concentrations of TRAP1-3 scaffold. The binding constant (K_D) of NADH respect to TRAP1-3 scaffold was 34.87 μM . The corresponding binding % of NADH to TRAP1-3 scaffold was fit to One Site-Specific binding model: $y = 132.6 \cdot x / (34.87 + x)$ with an R-squared of 0.9158.

In the light of these results, the scaffold physically brings the two dehydrogenases together and sequesters NAD cofactors in their surroundings at the concentration herein studied. When the reactions are carried out with 8-fold excess of free FDH1 (1.44 μM) to ameliorate the limitation this enzyme poses to the system, the positive effect of the scaffolding is dramatically reduced (Figure 5.14) In this scenario the higher FDH1 concentration reduces the average distance between enzymes to 104 nm as calculated with equation 1. Using an excess of FDH1 in solution both scaffolded and free systems performed the synthesis of L-Alanine 2 times faster (11.12 $\text{g} \times \text{g}_{\text{enzyme}}^{-1} \times \text{h}^{-1}$) and yielded higher product titers (75 mM) in 24 hours than using stoichiometric amounts of both dehydrogenases scaffolded in the TRAP1-3. Thus, the positive effect of the scaffolding on the system performance was minimized under these conditions. Overall, these data indicate that NAD species may be channeled between FDH1 and AlaDH3, thus enhancing the mass transfer of the cofactors between the two dehydrogenases.

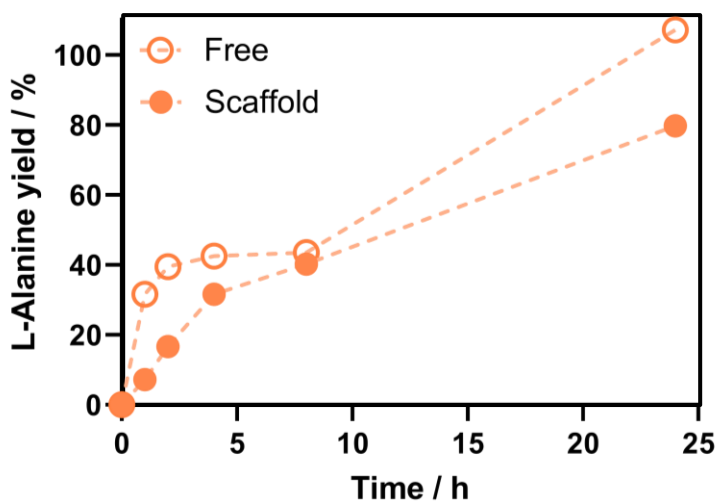


Figure 5.14. L-Alanine synthesis in batch-mode. L-Alanine yield at 8:1 enzyme ratio of FDH1:AlaDH3 for the free enzyme system (FDH1/AlaDH3) and for the scaffolded enzyme system (FDH1/AlaDH3@TRAP). Reaction mixture: 100 mM formate, 75 mM pyruvate, 500 mM ammonium chloride, 0.5 mM NADH and 0.15 mM FAD^+ .

5.2.4. Characterization of the channeling effect manifested by the scaffolded multi-enzyme systems

To experimentally support the channeling of NADH between the two dehydrogenases, we performed a competition assay with an NADH-dependent enzyme, and an isotope enrichment assay using deuterated formate. For the competition assay, we performed the reductive amination of pyruvate to L-Alanine in presence of an excess of free NADH oxidase (NOX) from *Thermus thermophilus* HB27 which oxidizes NADH with the concomitant production of hydrogen peroxide (H_2O_2). The H_2O_2 generated in situ can be

measured in line by adding a fourth enzyme, horseradish peroxidase (HRP), and Amplex red (AR) to the assay. If NADH reaches the active site of NOX, this enzyme produces H_2O_2 that is further used by HRP to transform AR into the fluorescent product resorufin (RSF) (Figure 5.15a). Through UV-vis quantification of the produced resorufin, we calculated the H_2O_2 generated and consequently, the NADH consumed, by NOX^{111,144}. Hence, the more active the NOX, the less efficient the NADH channeling between FDH1 and AlaDH3. First, we showed that AlaDH3 and NOX compete for NADH cofactor as NADH is consumed faster and H_2O_2 is formed when both enzymes are mixed. The mixture of HRP and AR with the AlaDH3 enzyme assay negligibly generated red color, confirming that H_2O_2 could only be formed by NOX. Then, the scaffolded and non-scaffolded systems were incubated with free NOX and HRP in presence of pyruvate and formate. Figure 5.15b shows that the NADH consumption was higher with the non-scaffolded system than with the scaffolded one. In agreement, the production of H_2O_2 was lower when the enzymes were scaffolded than when they were not. The H_2O_2 formation is the consequence of those NADH molecules that derail from the scaffolded cascade and escape to the bulk where are oxidized by NOX. Likewise, higher consumption of NADH means that FDH1 is not efficiently replenishing the pool of this reduced cofactor. Therefore, when FDH1 and AlaDH3 are scaffolded, NOX competes with AlaDH3 to a lower extent than when they are separated and diluted in the bulk. The competition assay informs us about the restricted leakage of NADH from the microenvironment of the scaffolded enzymes, supporting the fact that the cofactor regeneration, and thus the overall performance of the cascade, are more efficient when the enzymes co-localize in the TRAP-scaffold.

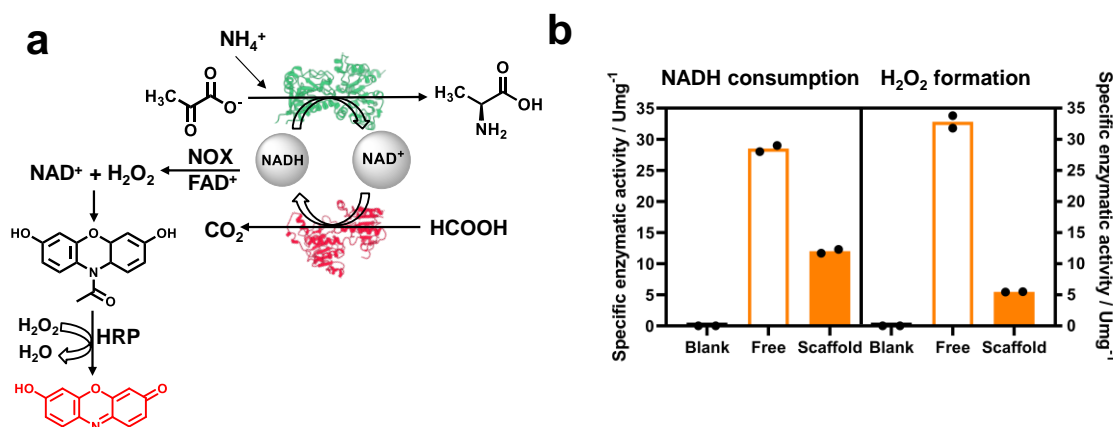


Figure 5.15. Side reaction competitive catalytic assay. a) Schematic representation of the competitive catalytic assay reaction. b) Competitive catalytic assay results reported as the consumption of NADH cofactor measured at 340 nm and the formation of H_2O_2 by measuring the production of resorufin at 560 nm. Both reactions using separately scaffolded enzymes (FDH1@TRAP/AlaDH3@TRAP, empty bars) or scaffolded enzymes in the same scaffold (FDH1/AlaDH3@TRAP, filled bars) were analyzed at both wavelengths to observe changes in the NADH consumption and H_2O_2 formation. Reaction mixture: 100 mM formate, 75 mM pyruvate, 500 mM ammonium chloride, 0.5 mM NADH, 0.15 mM FAD^+ , 0.1 mg/mL HRP, and 0.05 mM AR. The data are presented as the mean of two replicate experiments ($n = 2$).

To confirm the channeling of NADH between the two dehydrogenases, we performed a more complex experiment where an isotopic enrichment assay was merged with the competition assay previously mentioned. To do so, we performed the reaction under the presence of deuterated formate (see experimental section). In this assay, FDH1 produces deuterated NADH that is concurrently used by the neighbor AlaDH3 when they are scaffolded yielding deuterated L-Alanine-2-d. If the system is not scaffolded the deuterated NADH is released by FDH1 to the media where NOX can oxidize it back to NAD⁺ (Figure 5.17a). Under these conditions deuterated NADH could transfer the deuterium to pyruvate only if NADH is directly channeled from FDH1 to AlaDH3 (see methods). Unfortunately, the deuterated formate completely inhibited the FDH1 activity (Figure 5.16), thus the experiment was performed with a mixture of deuterated and non-deuterated formate (75 mM and 25 mM).

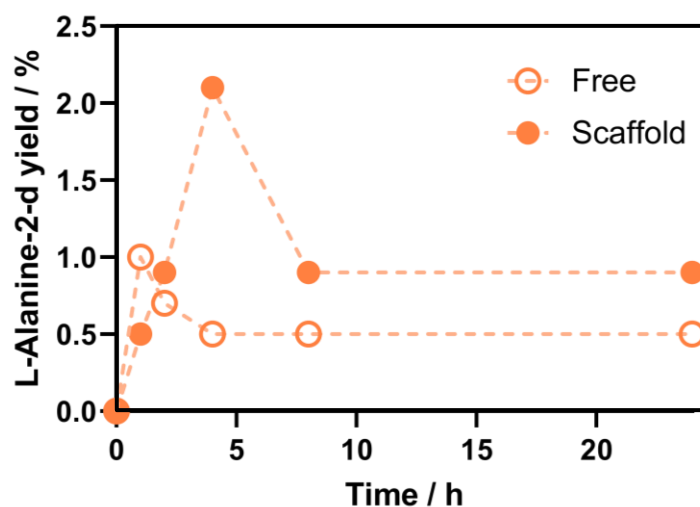


Figure 5.16. Deuterated L-Alanine synthesis in batch-mode. Deuterated L-Alanine yield % at 1:1 enzyme ratio of FDH1:AlaDH3 for the free enzyme system (FDH1/AlaDH3) and for the scaffolded enzyme system (FDH1/AlaDH3@TRAP). Reaction mixture: 100 mM deuterated formate, 75 mM pyruvate, 500 mM ammonium chloride, 0.5 mM NADH and 0.15 mM FAD⁺.

After 24 hours, 40 % of the L-Alanine synthesized was deuterated for the scaffolded system, whereas deuterated L-Alanine was undetected for the non-scaffolded system (Figure 5.17b). This enrichment in the heavy isotope of L-Alanine demonstrates that as soon as NADH is deuterated by FDH1 assembled on TRAP1-3, the neighboring enzyme, AlaDH3, uses this channeled cofactor to synthesize the deuterated L-Alanine. This experiment confirms the channeling of the cofactor between the two scaffolded dehydrogenases through a diffusion limited effect created by the postulated electrostatic interactions between the nicotinamide cofactors and the TRAP scaffold. This cofactor channeling contributes to speed up the cascade reaction and increases the final product yield.

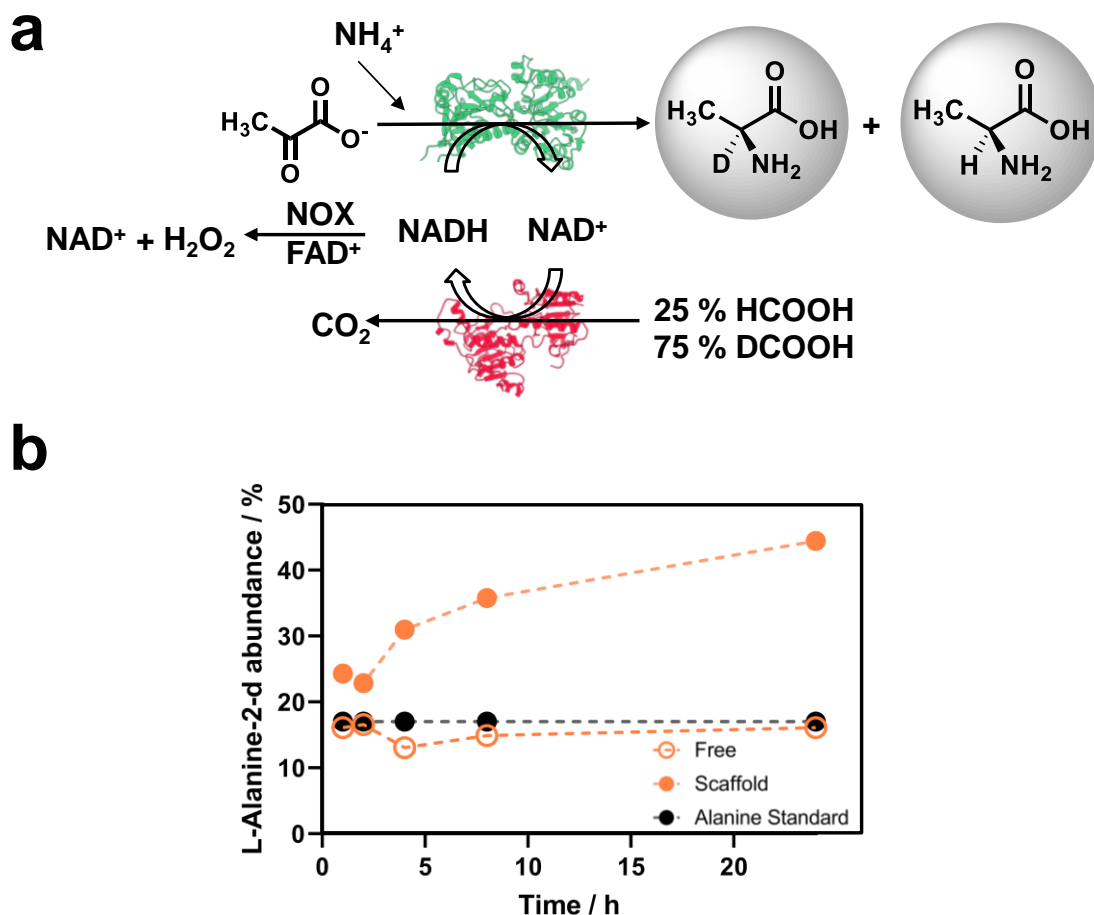


Figure 5.17. Isotopic enrichment assay coupled to competition assay for the synthesis of L-Alanine-2-d. a) Scheme of the catalytic cycle that was tested by the HPLC-MS. b) Isotopic abundance of deuterated L-Alanine-2-d ($M/z = 324$) for the free enzyme system (FDH1/AlaDH3, empty circles) and the scaffolded enzyme system (FDH1/AlaDH3@TRAP, filled orange circles). Reaction mixture: 75 mM deuterated formate, 25 mM formate, 75 mM pyruvate, 500 mM ammonium chloride, 0.5 mM NADH and 0.15 mM FAD^+ . Alanine standard is shown in filled black circles.

5.2.5. Heterogenization of the enzyme scaffolds

Once proved that the scaffolded system enhances the production rate and yield of the biosynthesis of L-Alanine through scaffold-assisted channeling of NADH, we intended to assemble the multi-enzyme system in solid-phase. Unfortunately, the solid-phase assembly failed to assemble the system with the optimal stoichiometry 6:3:1 (Table 5.6 and Figure 5.18). As model solid support for the scaffold immobilization, we selected agarose porous microbeads functionalized with cobalt-chelates on which the His-tag scaffoldin TRAP1-3 (His-Tagg-TRAP1-3) is selectively bound. First, we immobilized the His-TRAP1-3 on the agarose microbeads, achieving a protein load of $6.5 \text{ mg}_{\text{His-TRAP1,3}} \times \text{g}_{\text{carrier}}^{-1}$, which means a concentration of scaffold per volume of solid support of 0.14 mM. Then, we tested if agarose microbeads primed with the His-TRAP1-3 were capable of orderly assembling FDH1 and AlaDH3 on the solid phase. We follow a stepwise strategy similar to the one recently reported for the solid-phase assembly of cellulosome-based

enzymatic scaffold ³². Through mass balance between the offered solution and the supernatant upon the immobilization process, we calculated the load and the immobilization yield of each assembled enzyme. To quantify these parameters more accurately, we labeled the enzymes with fluorescent probes as follows, FDH1 with AF647, and AlaDH3 with AF488. Spectrophotometric quantification of the supernatants upon vacuum filtration informed us about the non-immobilized fraction of each enzyme, so we could indirectly calculate their bound fractions. In summary, the load of assembled FDH1 and AlaDH3 per mass of carrier was 4.52 mg x g⁻¹ and 4.45 mg x g⁻¹, respectively (Table 5.6).

Table 5.6. Immobilization parameters of sequentially co-immobilized FDH1/AlaDH3@TRAP on cobalt agarose (AG-Co²⁺) and directly co-immobilized FDH1/AlaDH3@TRAP and FDH1/AlaDH3 on tri-functional carriers. Amount of protein added to 1 g of the two carriers for the immobilization process. Amount of protein loaded on 1 g of the two carriers after the immobilization process. Immobilization yield, Ψ = (added protein/loaded protein on the different carriers) x 100.

Assembly element	Added protein / mg/g	Loaded protein on AG-Co ²⁺ carrier / mg/g	Ψ / %
TRAP1-3	6.5	6.5	100
FDH1	4.52	1.54	34
AlaDH3	4.45	1.51	34
Enzyme system	Added protein / mg/g	Loaded protein on tri-functional carrier / mg/g	Ψ / %
FDH1/AlaDH3@TRAP1-3	4.52/4.45@6.5	2.85/2.80@4.09	63
FDH1/AlaDH3	4.52/4.45	2.85/2.80	75

Accounting for loads of each element forming the scaffold, we obtained a scaffold with a molar ratio of 1:0.34:0.34 monomer-based, and 18:3:1 oligomer-based, for TRAP1-3:FDH1:AlaDH3. Hence, solid-phase assembly is less efficient than its counterpart in solution, which presented a molar ratio of 1:1:1. This result indicates that when two enzymes are assembled on an immobilized scaffold, the assembly is impaired by potential steric hindrances. The impaired assembly of FDH1 and AlaDH3 in solid phase was supported by lack of submicrometric co-localization found in the confocal laser scanning microscopy (CLSM) analysis (Figure 5.18) ¹¹³.

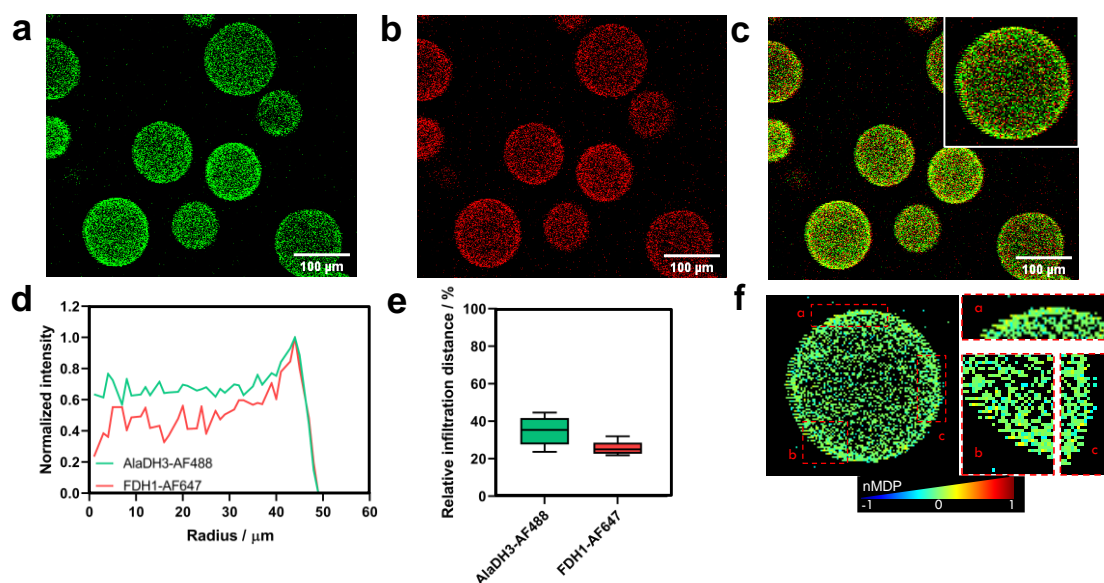


Figure 5.18. Confocal fluorescence microscopy images (20X magnification) of co-immobilized FDH1/AlaDH3@TRAP on AG-Co²⁺ carrier. AlaDH3 and FDH1 were labeled with AF488 and AF647. a) Spatial distribution of AlaDH3 (green channel), b) FDH1 (red channel) and c) overlay of the two fluorophores. The inset shows a digital zoom of the micrograph of c. d) Radial profile and e) relative infiltration distance of labeled FDH1 and AlaDH3. Relative infiltration distance is defined as the fraction of the radius where the fluorescence intensity was higher than 50 % of the maximum intensity. The data are presented as the mean of six replicate experiments (n = 6), and error bars represent standard deviations. f) Co-localization map created with Colormap Image J plugin. The scale refers to the co-localization degree of the two labelled enzymes. Blue pixels (value = -1) mean the absence of co-localization. Red pixels (value = 1) mean a high degree of co-localization. The regions framed with the dashed red square represent those pixels with co-localization values > 0.1 where the scaffolded has been assembled as the spatial colocalization of the two fluorophores is statistically significant.

To guarantee these ideal TRAP1-3:FDH1:AlaDH3 stoichiometry in the solid surface of the support, we first assembled the scaffold in solution and then immobilized it on porous agarose beads. Herein, the optimal stoichiometry of the scaffold in solution assures that both enzymes are close each other even when they are immobilized. As carrier, we exploited a tri-functional agarose-based carrier recently developed, which displays cobalt chelates, positively charged amine groups and aldehydes at its surface³³. The cobalt chelates drive the immobilization of the enzyme assembly through the His-tag fused to the TRAP1-3 unit, while aldehydes and amines establish covalent and ionic bonds with the enzyme complexes, respectively (Figure 5.19).

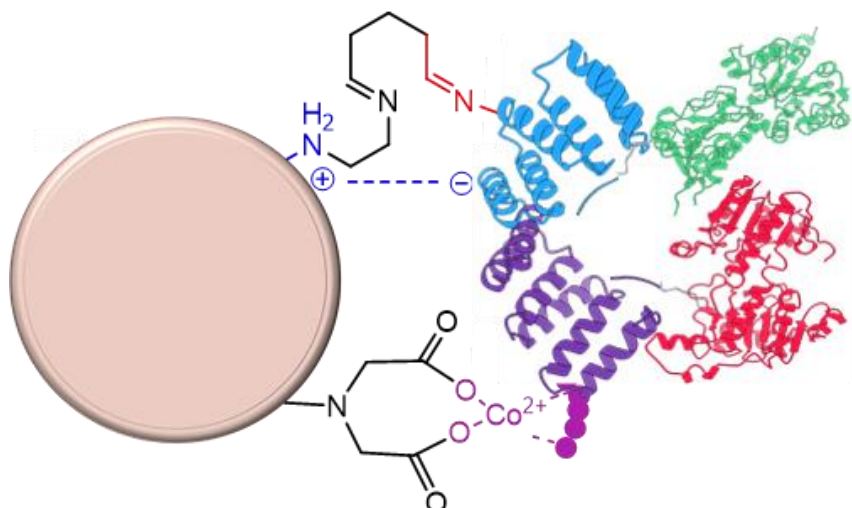


Figure 5.19. Scheme of the tri-functional carrier. The tri functionality is due to the presence of cobalt chelates, positively charged amine groups and aldehydes at its surface.

The load of FDH1/AlaDH3@TRAP renders enzyme loads of $2.8 \text{ mg} \times g_{\text{carrier}}^{-1}$ for AlaDH3 and FDH1 (Table 5.6). CLSM analysis reveals that the two enzymes are co-immobilized, but also co-localize on the same particle (Figure 5.20a-c) as their radial profiles and relative infiltration distances across the porous structure of the beads perfectly match (Figure 5.20d-e).

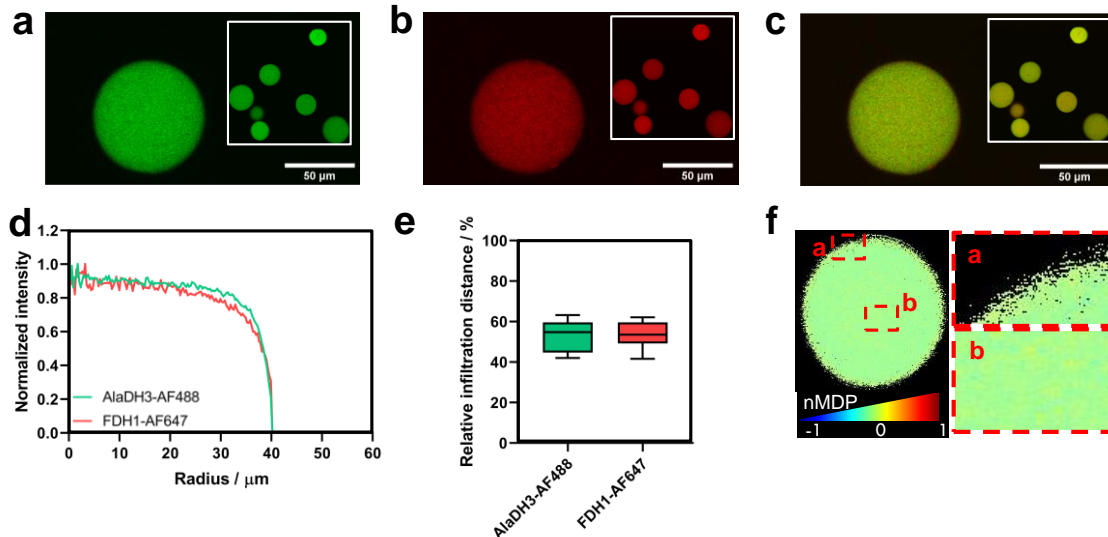


Figure 5.20. Confocal fluorescence microscopy images (20X magnification) of co-immobilized FDH1/AlaDH3@TRAP on tri-functional carrier. AlaDH3 and FDH1 were labeled with AF488 and AF647. a) Spatial distribution of AlaDH3 (green channel), b) FDH1 (red channel) and c) overlay of the two fluorophores. The inset shows a digital zoom of the micrographs. d) Radial profile and e) relative infiltration distance of labeled AlaDH3 (green line and bar) and FDH1 (red line and bar). Relative infiltration distance is defined as the fraction of the radius where the fluorescence intensity was higher than 50 % of the maximum intensity. The data are presented as the mean of ten replicate experiments ($n = 10$), and error bars represent standard deviations. f) Co-localization

map created with Colormap Image J plugin. The scale refers to the co-localization degree of the two labelled enzymes. Blue pixels (value = -1) mean the absence of co-localization. Red pixels (value = 1) mean a high degree of co-localization. The regions framed with the dashed red square represent those pixels with co-localization values > 0.1 where the spatial colocalization of the two fluorophores is statistically significant.

Furthermore, the colocalization analysis support that both enzymes greatly colocalize (Person coefficient of 0.85 ± 0.04 and Manders coefficients > 0.99) (Table 5.7). Moreover, the index of correlation obtained through colocalization map analyses presented similar conclusions (0.58 ± 0.04)^{145,146}. To assess the degree of co-localization between fluorophores, the Pearson correlation coefficient (PCC) and Manders overlap coefficient (MOC) are utilized. The two coefficients are mathematically equivalent, but they differ in their usage of either absolute intensities (MOC) or departure from the mean (PCC)¹⁴⁷. In confocal microscopy, the index correlation is the correlation of fluorescence intensity between spatially correlated enzymes immobilized in agarose beads¹⁴⁸.

Figure 5.20f shows higher values of the normalized mean deviation product (nMDP) in the outer surface (yellow pixels) and in the large voids of the agarose beads, pointing out that the fully assembled scaffold is mainly localized in highly open porous regions where steric hindrances are minimized. In parallel, we co-immobilized FDH1 and AlaDH3 without His-tag on the same carrier to fabricate a benchmarked heterogeneous biocatalyst that allowed us studying the effect of the spatial enzyme arrangement in the confined space. In this latter case, the two enzymes were randomly immobilized through electrostatic and covalent interactions established between acidic residues (Asp and Glu) and Lys of the enzymes and the amine and aldehyde groups of the carrier, respectively. The enzyme load was similar to the scaffolded system ($3.3 \text{ mg} \times \text{g}_{\text{carrier}}^{-1}$ for both enzymes, Table 5.6), maintaining the same AlaDH3:FDH1 monomer ratio. In this case, the two enzymes co-localized to a lower extent than when the scaffolded system is immobilized but to a larger extent than when the scaffold is assembled in solid phase (Figure 5.21 and Table 5.7).

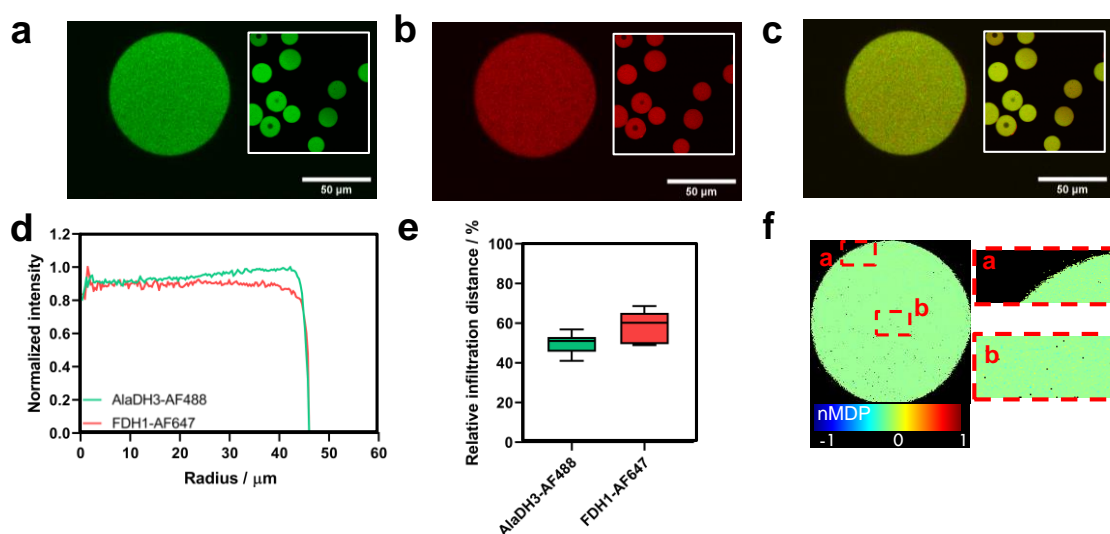


Figure 5.21. Confocal fluorescence microscopy images (20X magnification) of co-immobilized FDH1/AlaDH3 on tri-functional carrier. AlaDH3 and FDH1 were labeled

with AF488 and AF647. a) Spatial distribution of AlaDH3 (green channel), b) FDH1 (red channel) and c) overlay of the two fluorophores. The inset shows a digital zoom of the micrographs. d) Radial profile and e) relative infiltration distance of labeled AlaDH3 and FDH1. Relative infiltration distance is defined as the fraction of the radius where the fluorescence intensity was higher than 50 % of the maximum intensity. The data are presented as the mean of ten replicate experiments ($n = 10$), and error bars represent standard deviations. f) Co-localization map created with Colormap Image J plugin. The scale refers to the co-localization degree of the two labelled enzymes. Blue pixels (value = -1) mean the absence of co-localization. Red pixels (value = 1) mean a high degree of co-localization. The regions framed with the dashed red square represent those pixels with co-localization values > 0.1 where the spatial colocalization of the two fluorophores is statistically significant.

Table 5.7. Pearson Manders and co-localization coefficient determined through the analysis of confocal microscopy images displayed with FIJI software using JaCoP and co-localization colormap plugins. Pearson coefficient expresses the intensity correlation of fluorescence in two images. Manders coefficient A (MA) informs about the proportion of AF647-labeled FDH1 that overlaps in the space with the AF488-labeled AlaDH3, while Mander coefficient B (MB) reflects the opposite, *i.e.* the proportion of AF488-labeled AlaDH3 that overlaps in the space with the AF647-labeled FDH1. Index of correlation informs about the same as Pearson coefficient, indicating the distribution of the regions where pixels of both enzymes are present. The standard deviation was calculated based on the measurements on the 5 beads. System 1 is the solid-phase assembly of the scaffold using His-TRAP1-3 as priming unit previously immobilized on AG-Co²⁺. System 2 is the scaffold assembled in solution and subsequently immobilized on the trifunctional agarose-based carrier activated with aldehydes, cobalt chelates and positively charged amine groups. System 3 is the benchmarked system where non-scaffolded AlaDH3 and FDH1 are immobilized on the tri-functional carrier.

System	Coef. Pearson	Coef. MA	Coef. MB	Index of correlation
1	0.29 ± 0.05	0.39 ± 0.03	0.14 ± 0.02	0.40 ± 0.03
2	0.82 ± 0.05	1 ± 0	0.99 ± 0	0.58 ± 0.04
3	0.85 ± 0.04	0.99 ± 0	0.93 ± 0.07	0.55 ± 0.04

Once the scaffolded and non-scaffolded systems were prepared and characterized at the sub-micrometric level, they were challenged for the reductive amination of pyruvate with *in situ* NADH recycling using formic acid as ancillary electron donor³⁴. We assessed the performance of the different spatial arrangements both in solution and immobilized. In all cases, we adjusted the molar ratio of the four systems to 1:1 FDH1:AlaDH3. Reactions were carried out as described above and the L-Alanine yield was determined at 24 hours by HPLC. Figure 5.22a shows that L-Alanine yield decreases two times when the scaffolded system was immobilized compared to its counterpart in solution. Likewise, the specific productivity of the scaffolded system in solution was 5 times higher than its immobilized counterpart (Table 5.8).

Table 5.8. Specific productivity of L-Alanine ($\text{g} \times \text{g}_{\text{enzyme}}^{-1} \times \text{h}^{-1}$) for the biotransformation catalyzed by the soluble and immobilized systems.

Formulation	FDH1/AlaDH3	FDH1/AlaDH3@TRAP
Soluble	0.92	5.21
Immobilized	0.98	1.11

The lower yield and productivity measured upon immobilization might be explained by both external and internal diffusion limitations typically observed for heterogeneous biocatalysts. To note, the scaffolded multi-enzyme system achieved roughly two times larger L-Alanine chromatographic yields regardless of whether it was immobilized or free. Then we recycled the two scaffolded and non-scaffolded immobilized systems for consecutive operational cycles. Upon each 24 hours reaction cycle, the solid biocatalysts were separated from the reaction media through vacuum filtration and washed before starting a new cycle. The reaction crudes were analyzed by HPLC to determine the L-Alanine yield. Figure 5.22b shows a decay of the product yield along the cycles indicating the inactivation of the heterogeneous biocatalysts. The immobilized and scaffolded system was more stable than its non-scaffolded counterpart as the former reached 18 % product yield after the third cycle, compared to the 10 % yield achieved by the latter. Despite none of the tested immobilized systems achieved 100 % product yield and were relatively unstable during their operational use under the tested conditions, the scaffolding of the two enzymes overall enhances the productivity and the operational stability of the multi-enzyme system.

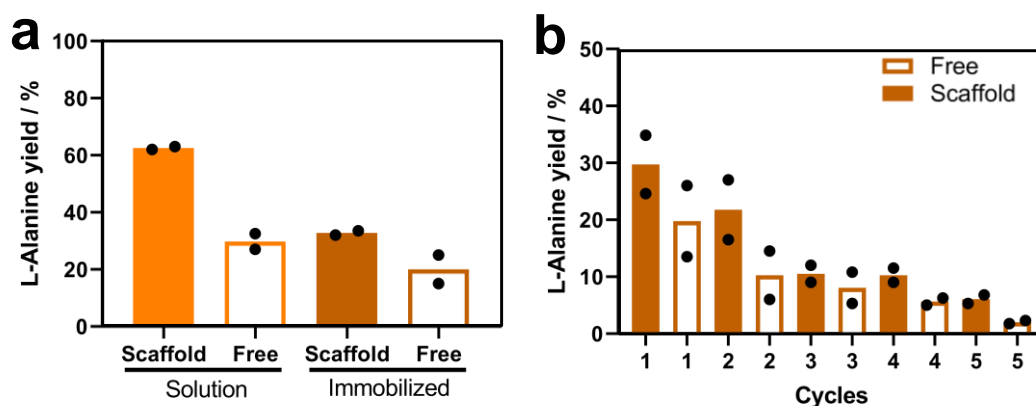


Figure 5.22. Operational performance and reusability of the scaffolded enzyme systems immobilized on porous carriers. a) L-Alanine yield % comparing different multi-enzyme systems in solution and immobilized after 24 hours. b) L-Alanine yield % in a reusability test for free (FDH1/AlaDH3) (empty bars) and scaffolded (FDH1/AlaDH3@TRAP) (orange bars) enzyme systems. Reaction mixture: 100 mM formate, 75 mM pyruvate, 500 mM ammonium chloride and 0.5 mM NADH. The data are presented as the mean of two replicate experiments ($n = 2$).

To understand if the biocatalysts inactivation was due to either enzyme leakage or intrinsic enzyme inactivation, we analyzed the reaction crudes and the washes by SDS-PAGE. Since both immobilization chemistries for scaffolded and non-scaffolded systems rely on a covalent immobilization chemistry, leaching of the working enzymes should

only occur if some enzyme subunits do not directly interact with the carrier surface. SDS-PAGE analysis negligibly detected proteins in the reaction crude and wash solution after each cycle when the enzyme system was scaffolded and immobilized, on the contrary when the two enzymes were co-immobilized but not scaffolded, we can detect the lixiviation of some AlaDH subunits (Figure 5.23) after two reaction cycles. Therefore, we suggest that the decay in the performance of the co-immobilized but not scaffolded system is practically due to the dissemble of the quaternary structure of AlaDH which is stabilized to a higher extent when it is scaffolded. Nonetheless, we cannot discard that beside lixiviation, the activity loss of both systems is also driven by inactivating structural distortions that occur during their operational use. Therefore, product yield and SDS-analysis support the fact that the scaffolding of the system promotes a significant overall operational stabilization of the assembled multi-enzyme system.

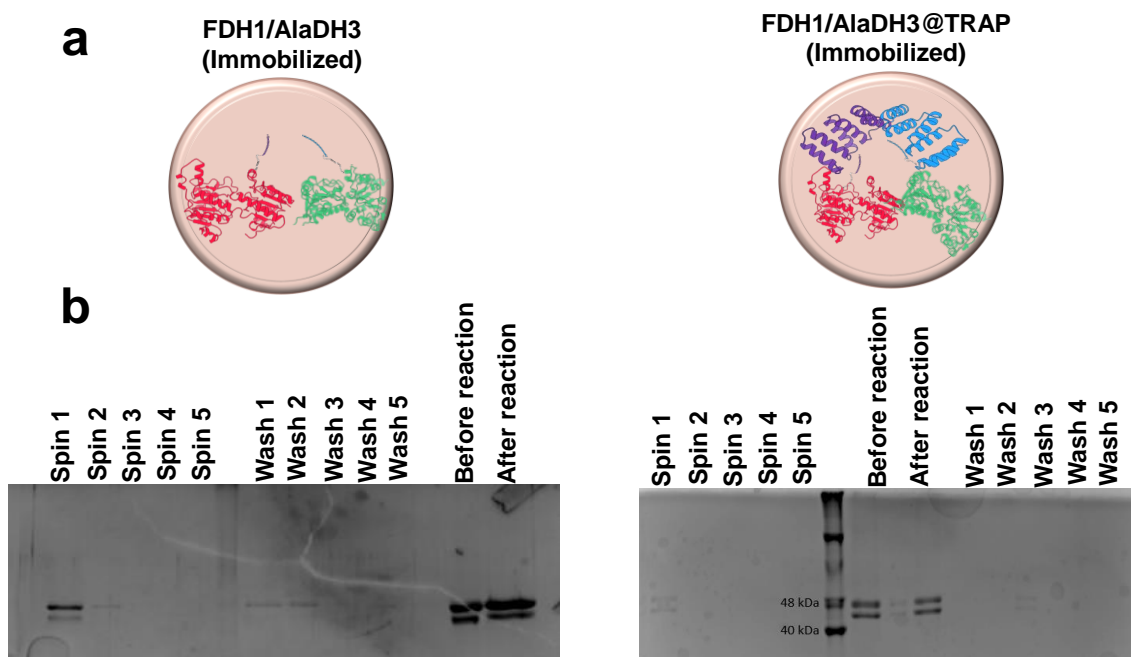


Figure 5.23. Immobilization of free and scaffolded enzyme systems. Immobilization of free and scaffolded FDH1/AlaDH3 enzyme systems. a) Scheme of the immobilized systems analyzed in the assay: FDH1/AlaDH3 and FDH1/AlaDH3@TRAP. b) SDS-PAGE gel electrophoresis of FDH1 (44 KDa band) and AlaDH3 (48 KDa band) enzymes before and after performed the reusability test. Spin samples are the reaction crude after 24h. Wash samples were the wash of the heterogeneous biocatalysts after each cycle with 10 volumes of reaction buffer. The spin and wash steps were shown in order to follow the process. Left panel: reusability test of free enzyme system, FDH1/AlaDH3. Right panel: reusability test of scaffolded enzyme system, FDH1/AlaDH3@TRAP.

5.2.6. Amine biosynthesis catalyzed by a TRAP-scaffolded tri-enzyme system

Encouraged by the enhancement that the TRAP-based scaffold promoted in the redox FDH1/AlaDH3 system for the synthesis of L-Alanine, we went one step further and assembled a tri-enzyme system onto the TRAP scaffold⁹³. As third enzyme we selected the ω -transaminase from *Pseudomonas fluorescens* (ω TA). The aim of incorporating this new enzyme was to construct a cascade with three orthogonal reactions to *in situ* recycle both redox cofactors and amine donors. This cascade has been previously designed to

aminate benzaldehyde to benzylamine using L-Alanine as amine donor (Figure 5.24a)¹⁴⁹. This amine donor is *in situ* regenerated from the pyruvate released by the ω TA, ammonium formate and NADH through a reductive amination catalyzed by AlaDH. Finally, the NADH pool is replenished by the action of FDH which uses formate as hydride donor to reduce the NAD⁺ to NADH (Figure 5.24a). Thus, the couple AlaDH3/FDH1 allows the recycling of both the amine donor and the redox cofactor^{149,150}. This cascade is thermodynamically challenging as concentration-based equilibrium constant of the amine donor (L-Alanine) and the molar ratio between the amine donor and amine acceptor (benzylaldehyde) are extremely low from the beginning of the reaction^{151,152}.

To assemble this tri-enzyme system we constructed a new TRAP-scaffold with an additional binding domain (TRAP2) for anchoring ω TA. TRAP2 domain has high affinity ($K_D = 1.7 \mu\text{M}$) and selectivity toward the peptide-2 MERVW, and is orthogonal to peptide-1 and peptide-3 displaying weaker affinity with K_D values of 42.1 μM and 83 μM , respectively⁹³. In parallel, this peptide-2 was fused to the C-terminus of ω TA giving rise to the variant ω TA2. Before carrying out the assembly, the structural integrity and functionality of ω TA2 and TRAP2-3-1 were characterized by SDS-PAGE, MALDI-TOF and CD, confirming the expected size and secondary structure according to their primary sequences (Table 5.1). The new trivalent scaffold (TRAP2-3-1) and the corresponding tagged enzymes were assembled following the stepwise protocol described above. The assembly sequence was AlaDH3, FDH1, and ω TA2 to achieve the tri-enzyme scaffold. SEC analysis showed a main peak that corresponded to a molecular size of 605 kDa, which indicates that the scaffold was also assembled in a equimolar ratio 1:1:1:1 of TRAP2-3-1:FDH1: ω TA2:AlaDH3 of the monomeric units, according to the primary sequence of all elements. In addition, when this SEC peak was analyzed by SDS-PAGE gel electrophoresis, we could identify the three enzymes and the scaffold (Figure 5.6c).

After assembly, the activity of ω TA2 was spectrophotometrically analyzed by the broadly used transaminase assay based on the deamination of methylbenzylamine (FEA)¹⁵³. Figure 5.24b shows that the scaffolded ω TA2 exhibited 74 % of the activity of its free counterpart, considering that the specific enzyme activity of the ω TA enzyme was significantly affected after fusion to peptide-2 with a 42 % decrease in activity (Table 5.3). Similarly, to FDH1 and AlaDH3, the reduction of the enzyme activity upon the assembly might be owing to steric hindrances the small substrates (L-Alanine and benzaldehyde) suffer to reach the more crowded environment of scaffolded ω TA2. When the scaffold FDH1/ ω TA2/AlaDH3 was incubated with benzaldehyde, an excess of pyruvate and ammonium, and substoichiometric amounts of NADH, we observed that the benzylamine chromatographic yield (CY) upon 24 hours was 50 % when using the scaffolded system; a yield 4.25-fold larger than using the free system (Figure 5.24c). While the scaffold system reached a plateau after 24 h achieving a titer of 5.1 mM benzylamine, the free system keeps working until reaching a titer of 5.9 mM benzylamine after 48 hours. Likewise, the system productivity was enhanced 2.36 times when the three enzymes were scaffolded compared to the free system after 48 h. Likewise, the system productivity was enhanced 2.36 times when the three enzymes were scaffolded compared to the free system after 48 h. The maximum specific productivity of the scaffolded biocatalyst was $0.047 \text{ g} \times \text{g}_{\text{enzyme}}^{-1} \times \text{h}^{-1}$, a value that is 8-times larger than the reductive aminating cascade composed by ω TA from *Aspergillus terreus*, AlaDH, and

FDH reported elsewhere¹⁵⁴. The scaffolded system reaches the same specific productivity as ω TA cross-linked enzyme aggregates (CLEA) working as sole enzyme in the reductive amination of benzaldehyde using a 10 times molar excess of L-Alanine as amine donor ($0.049 \text{ g} \times \text{g}_{\text{enzyme}}^{-1} \times \text{h}^{-1}$)⁹⁷. Despite the formation of benzylamine in presence of an excess of pyruvate is thermodynamically unfavoured, the physical proximity of the amine donor generating enzymes (FDH and AlaDH) and the ω TA afford an amination reaction otherwise limited using a diluted free enzyme system. This scaffolded system with three enzymes is one of the few examples of assemblies with more than two enzymes using biomolecular scaffolds that do not rely on enzyme clustering. TRAP2-3-1 is able to form a supramolecular complex that gathers two dehydrogenases and one transaminase to force an inter-enzyme channeling from one enzyme to its assembled neighbor in a thermodynamically unfavored cascade. As well as in the case of the bi-enzyme system FDH1/AlaDH3, we suggest a channeling driven by NADH-TRAP2-3-1 interactions.

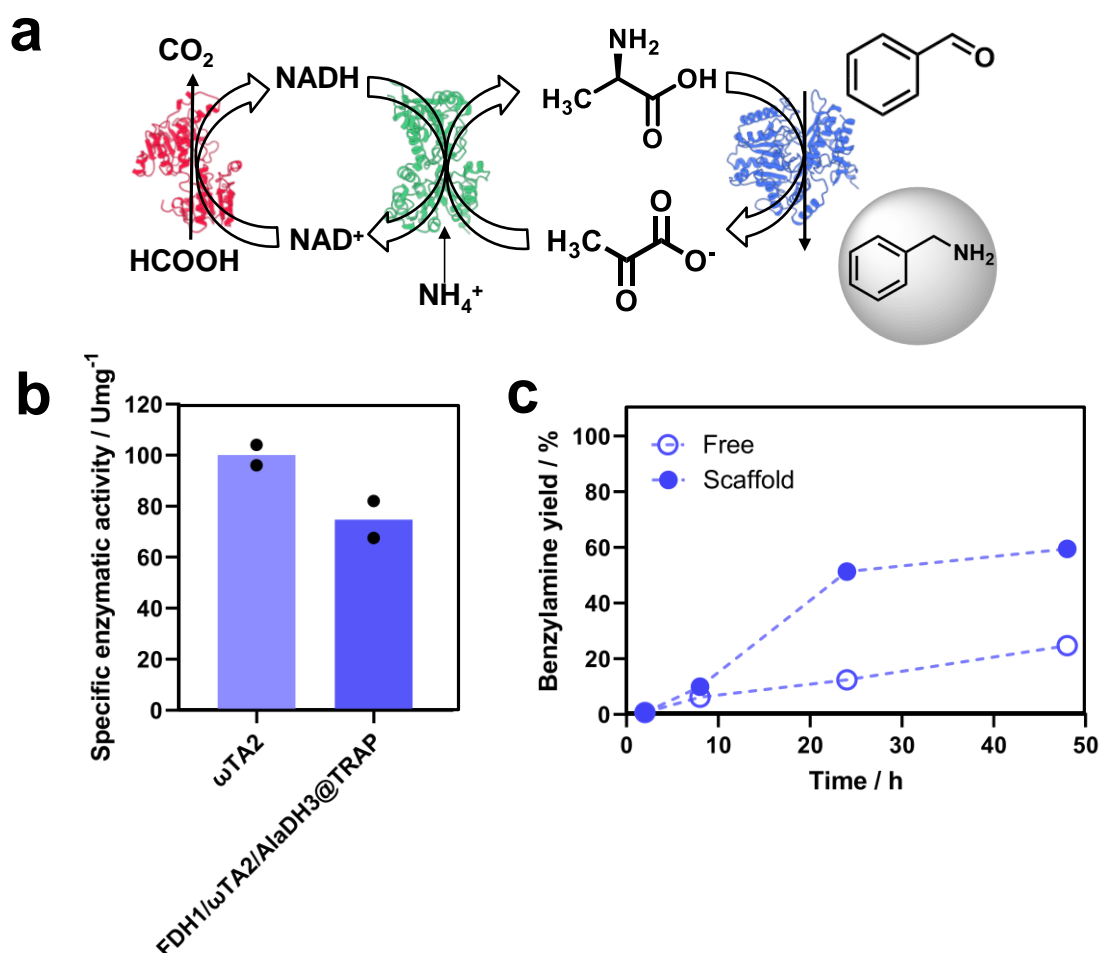


Figure 5.24. Benzylamine biosynthesis using a three-enzyme organized multi-enzymatic system. a) Scheme of the parallel reactions to produce benzylamine. b) Relative activity of ω TA2 in the free and scaffolded form (FDH1/ ω TA2/AlaDH3@TRAP). 100 % of ω TA activity corresponds to $5.58 \text{ U} \cdot \text{mg}^{-1}$. c) Benzylamine chromatographic yield % at an enzyme monomer ratio of 1:1:1 for in free (FDH1/ ω TA2/AlaDH3) and scaffolded system (FDH1/ ω TA2/AlaDH3@TRAP). Reaction mixture: 10 mM benzaldehyde, 50 mM

pyruvate, 500 mM ammonium formate, 0.1 mM PLP and 0.5 mM NADH. The data are presented as the mean of two replicate experiments (n = 2).

5.3. Conclusions

In conclusion, a simple methodology has been developed for assembling of enzymes onto a scaffold based on engineered Tetratricopeptide Repeat Affinity Proteins (TRAP). These modules can be coupled to form arrays of specific and orthogonal recognition protein domains that selectively bind short peptide sequences fused to enzymes. This is a methodology driven by a generally applicable biomolecular recognition which serves as a tool for preparing scaffolded multi-enzyme systems. In addition to the spatial organization achieved by this strategy, it should be noted that TRAP scaffolds can also be engineered to electrostatically interact with cofactors and reaction intermediates, providing a second level of tunability to the system. To the best of our knowledge, this is the first example of a protein-scaffold designed to organize several enzymes at the nanoscale and capture reaction intermediates to increase their local concentration in the surroundings of the scaffolded enzymes. This has been shown previously with DNA scaffold but never with protein-based ones.

These concepts have been proven with a two-enzyme system that performs the simultaneous asymmetric reduction of α -ketoacids to L-amino acids while recycling NADH. These enzymes were successfully assembled on the TRAP1-3 scaffold via biomolecular recognition with a controlled stoichiometry, where the largest enzyme acts as a nucleation point for the assembly. The efficiency of the biocatalytic cascade was significantly improved upon assembly due to a diffusion-limited strategy driven by electrostatic cofactor-scaffold interactions, resulting in the effective channeling of NADH from FDH1 to the adjacent AlaDH3 and boosting *in situ* NADH recycling. This channeling effect, albeit to a lesser extent, also occurred when the enzyme scaffold was immobilized on porous beads. Finally, a more complex system composed of three-tagged enzymes FDH1, ω TA2, and AlaDH3, was assembled on a scaffoldin unit with three orthogonal TRAP domains (TRAP2-3-1). This cascade with three parallel reactions allows *in situ* recycling of both redox cofactors and amine donors. The assembly of this three-enzyme cascade demonstrated not only that the methodology developed allows the coupling of three step cascade reactions, but also that the throughput for the cascade is enhanced by arranging the enzymes at the nanoscale and by increasing the local concentration of the cofactors by directed interactions with the scaffold.

The methodology developed here is relatively simple and modular compared to other current approaches. This is due the following facts: 1) the scaffolds are exquisitely orthogonal; 2) the peptide tagging approach is simple, as only short 10-20 amino acid sequences are needed to anchor enzymes to the scaffoldin; 3) the approach is easily genetically programmed, as we constructed a set of plasmids for block cloning and tagging of any enzyme; 4) the peptide tagging has not resulted in any major defect on the protein activity for all the systems tested to date; 5) the cofactor interactions are driven by easily encodable electrostatic interactions; and finally 6) the methodology also involves a straightforward incubation process for assembling the components, which makes it relatively simple and scalable for use in large-scale applications.

Chapter 5. Engineered repeat proteins as scaffolds to assemble multi-enzyme systems for efficient cell-free biosynthesis

We envision that this technology will make a strong contribution to advancing the manufacture of more robust multi-enzyme systems, where enzymes are scaffolded and organized with nanometric precision. Our endeavor has been based on creating a relevant multi-enzyme system that allows for the assembling of enzymes capable of acting as biocatalysts on protein-based scaffolds. Therefore, the combination of multifactor protein engineering and biocatalysis exhibits enormous potential not only to enhance the intrinsic catalytic activity and stability of enzymes but also to maximize the performance of spatially organized multi-enzyme systems¹⁵⁵. The controlled interactions with key cofactors and reaction intermediates make this technology even more promising for a wide range of applications in biocatalysis. Furthermore, the applications demonstrated in this work could be expanded to other fields of applied science, for instance, the integration into energy devices,¹⁵⁶ or the formation of biocatalytic films¹⁵⁷.

General conclusions

This thesis focused on designing synthetic proteins as scaffolding units for assembling multi-enzymatic systems.

The rational design of synthetic protein-based scaffolds plays a crucial role in organizing enzymes inside cells, facilitating substrate transport, and creating an optimal microenvironment for enzyme function. By combining protein engineering and biocatalysis, efficient multi-enzyme systems can be designed from the bottom up, utilizing nanoscale positioning to enhance catalytic performance. Therefore, this thesis explores how different parameters affect the catalytic efficiency of multi-enzyme scaffold systems in a systematic manner. Tetratricopeptide repeat (TPR) engineered protein scaffolds used for this study have modular and simple engineering features, allowing them to assemble as scaffold units and regulate interactions with critical cofactors and intermediates, as well as precise control over spatial arrangements at the nanoscale.

In this work, two scaffolding approaches were explored to achieve this objective: a supramolecular assembly of protein modules with intrinsic self-assembly properties and scaffolding based on biomolecular recognition.

- **In the supramolecular assembly approach, enzymes were directly fused to protein modules with intrinsic assembly properties.**
- **In the biomolecular recognition-based scaffolding approach, enzymes were fused to peptide tags for assemble into scaffolds with orthogonal peptide recognition properties.**

Using protein engineering and protein assembly approaches, we achieved multi-enzyme systems with precise control at the nanometer scale over the arrangement of the assembled enzymes. Finally, we developed two organized multi-enzyme systems and evaluated them in two industrially relevant biotransformations, leading to the production of the natural amino acid L-Alanine and the organic molecule benzylamine.

The conclusions can be summarized in four main points:

TPR modules can be successfully engineered to generate assemblies with orthogonality properties, and precise spatial control. In our study, we investigated TPR engineered modules, specifically CTPR and TRAP, as a foundation for the development of three protein-based scaffolding strategies.

Two of the scaffolding strategies explored the feasibility of utilizing genetically programmed intermolecular staples to direct the assembly of CTPR domains. To facilitate this process, we employed two distinct approaches to design intermolecular staples that possessed orthogonality and were capable of stabilizing assemblies formed through intrinsic head-to-tail interactions between CTPR modules. For our **first and second scaffolding strategies**, we implemented **cysteine-mediated di-sulfide bonds and metal-directed assembly** as stapling chemistries, effectively securing the assemblies. The overall design concept focused on encoding orthogonality into the CTPR systems, which already possessed self-assembly properties.

The **third scaffolding strategy** leveraged **biomolecular recognition** by employing TRAP modules with distinct orthogonal biorecognition sites within the same scaffold that interact

General conclusions

with their corresponding tag-peptides. These interactions facilitate the assembly of various macromolecules, including proteins, onto the TRAP scaffolds. By capitalizing on the unique recognition sites present within the TRAP modules, this strategy enables a selective and precise assembly process, enabling controlled attachment of specific molecules onto the scaffold. Furthermore, the short sequences encoding the tag peptides utilized in this approach are simply fused to the biomolecule of interest, ensuring that the properties of the selected biomolecules remain unaltered.

In summary, our study demonstrates the successful use of TPR engineered modules, specifically CTPR and TRAP, in three protein-based scaffolding strategies.

The engineered TPR synthetic protein scaffolds were demonstrated to effectively assemble enzymes. We employed CTPR modules, specifically engineered SCABs, to directly fuse selected dehydrogenase enzymes (FDH and AlaDH) and thus facilitate their assembly. For this purpose, we used the two SCAB-based assembly strategies developed and driven by cysteine stapling, and metal-directed assembly of the modules. SCAB-enzyme fusions were assembled successfully through both approaches. For the third scaffolding strategy, the chosen dehydrogenase enzymes (FDH and AlaDH) were fused to TRAP-tag peptides rather than directly to the TRAP proteins. The assembly process relied on a biomolecular recognition mechanism between the peptides fused to the enzymes and the corresponding recognition sites on the TRAP scaffold. Through this approach, we also achieved controlled and selective assembly, demonstrating the potential of utilizing biomolecular recognition for constructing complex and spatially organized enzyme systems.

Scaffolded enzymes retain their enzymatic activity within novel organized multi-enzyme systems. We focused first on the two scaffolded dehydrogenase enzymes (FDH-AlaDH) and compared their enzyme activity in three distinct scaffolded systems with that of the free enzymes. Importantly, we found that the scaffolded enzymes retained their enzyme activity even after the assembly processes. This corroborates the successful integration of the dehydrogenase enzymes into the organized multi-enzyme systems and highlights the potential of utilizing scaffolding strategies to preserve enzyme functionality in complex assemblies.

Scaffolding of multi-enzymatic systems on engineered TPR scaffolds resulted in highly efficient biocatalysts. Through the implementation of novel organized multi-enzyme scaffolded systems, we observed increased catalytic activity in two out of the three developed systems compared to free enzymes. This resulted in enhanced production of L-Alanine, a natural amino acid, and regeneration of the highly sought-after cofactor NADH. Notably, we expanded our efforts by constructing a multi-enzyme scaffolded system based on biomolecular recognition assembly, incorporating three enzymes (FDH, ω TA, and AlaDH). Within this organized tri-enzyme scaffolded system, we successfully demonstrated the synergistic cooperation between cascade reactions. The FDH-AlaDH pair effectively generated L-Alanine while regenerating NADH, while ω TA utilized this L-Alanine for the amination of benzaldehyde into benzylamine. Furthermore, the production of benzylamine in the organized tri-enzyme scaffolded system exceeded the efficiency of free enzymes. These findings underscore the importance of rational protein design, considering factors such as enzyme proximity, scaffold physicochemical properties, and the resulting enhanced catalytic performance observed in TPR protein-based scaffolded

systems. Additionally, the coding of interactions with products, substrates, or cofactors contributes to high local concentrations and spatial organization, further enhancing the overall functionality of the scaffolded systems.

Overall, this thesis introduces synthetic protein scaffolds for the controlled organization of multi-enzyme systems at the nanoscale, providing an adaptable and versatile technology for diverse biocatalytic applications. This research contributes to the advancement of knowledge in protein science, enabling the development of more robust and tunable protein scaffolds. These scaffolds can ultimately be applied to arrange different complex biomolecular systems, particularly multi-enzymatic systems. This technology is also expandable to other areas of applied science, such as integration into energy devices, or the formation of biocatalytic films, leading to the generation of highly efficient and precisely positioned multi-enzyme systems in various applied science fields.

References

1. Fasim, A., More, V. S. & More, S. S. Large-scale production of enzymes for biotechnology uses. *Curr. Opin. Biotechnol.* **69**, 68–76 (2021).
2. Abedi, D., Zhang, L., Pyne, M. & Chou, C. P. Enzyme biocatalysis. *Comprehensive Biotechnology.* **1**, 15-24 (2011).
3. Ellis, G. A. *et al.* Artificial Multienzyme Scaffolds: Pursuing in Vitro Substrate Channeling with an Overview of Current Progress. *ACS Catal.* **9**, 10812–10869 (2019).
4. Garcia-Galan, C., Berenguer-Murcia, Á., Fernandez-Lafuente, R. & Rodrigues, R. C. Potential of different enzyme immobilization strategies to improve enzyme performance. *Adv. Synth. Catal.* **353**, 2885–2904 (2011).
5. Yushkova, E. D. *et al.* Application of Immobilized Enzymes in Food Industry. *J. Agric. Food Chem.* **67**, 11553–11567 (2019).
6. Al-Ghanayem, A. A. & Joseph, B. Current prospective in using cold-active enzymes as eco-friendly detergent additive. *Appl. Microbiol. Biotechnol.* **104**, 2871–2882 (2020).
7. Kaushal, J., Mehandia, S., Singh, G., Raina, A. & Arya, S. K. Catalase enzyme: Application in bioremediation and food industry. *Biocatal. Agric. Biotechnol.* **16**, 192–199 (2018).
8. Fessner, W. D. Systems Biocatalysis: Development and engineering of cell-free ‘artificial metabolisms’ for preparative multi-enzymatic synthesis. *N. Biotechnol.* **32**, 658–664 (2015).
9. Schmidt-Dannert, C. & Lopez-Gallego, F. A roadmap for biocatalysis – functional and spatial orchestration of enzyme cascades. *Microb. Biotechnol.* **9**, 601–609 (2016).
10. Lopez-Gallego, F. & Schmidt-Dannert, C. Multi-enzymatic synthesis. *Curr. Opin. Chem. Biol.* **14**, 174–183 (2010).
11. Köhler, V. & Turner, N. J. Artificial concurrent catalytic processes involving enzymes. *Chem. Commun.* **51**, 450–464 (2015).
12. Rocha-Martín, J., Rivas, B. de Las, Muñoz, R., Guisán, J. M. & López-Gallego, F. Rational co-immobilization of bi-enzyme cascades on porous supports and their applications in bio-redox reactions with insitu recycling of soluble cofactors. *ChemCatChem* **4**, 1279–1288 (2012).
13. Velasco-Lozano, S., Benítez-Mateos, A. I. & López-Gallego, F. Co-immobilized Phosphorylated Cofactors and Enzymes as Self-Sufficient Heterogeneous Biocatalysts for Chemical Processes. *Angew. Chemie - Int. Ed.* **56**, 771–775 (2017).
14. Bolivar, J. M., Woodley, J. M. & Fernandez-Lafuente, R. Is enzyme immobilization a mature discipline? Some critical considerations to capitalize on the benefits of immobilization. *Chem. Soc. Rev.* **51**, 6251–6290 (2022).
15. Mateo, C., Palomo, J. M., Fernandez-Lorente, G., Guisan, J. M. & Fernandez-Lafuente, R. Improvement of enzyme activity, stability and selectivity via immobilization techniques. *Enzyme Microb. Technol.* **40**, 1451–1463 (2007).
16. Markel, U. *et al.* Advances in ultrahigh-throughput screening for directed enzyme evolution. *Chem. Soc. Rev.* **49**, 233–262 (2020).

References

17. Woodley, J. M. Enzyme Cascade Process Design and Modelling. *Enzyme Cascade Design and Modelling*. 125-139 (2021).
18. Chen, X., Zaro, J. L. & Shen, W. C. Fusion protein linkers: Property, design and functionality. *Adv. Drug Deliv. Rev.* **65**, 1357–1369 (2013).
19. Gokhale, R. S. & Khosla, C. Role of linkers in communication between protein modules. *Curr. Opin. Chem. Biol.* **4**, 22–27 (2000).
20. Wahab, R. A., Elias, N., Abdullah, F. & Ghoshal, S. K. On the taught new tricks of enzymes immobilization: An all-inclusive overview. *React. Funct. Polym.* **152**, 104613 (2020).
21. Liu, Q., Xun, G. & Feng, Y. The state-of-the-art strategies of protein engineering for enzyme stabilization. *Biotechnol. Adv.* **37**, 530–537 (2019).
22. Katchalski-Katzir, E. Immobilized enzymes - learning from past successes and failures. *Trends Biotechnol.* **11**, 471–478 (1993).
23. Arana-Peña, S. *et al.* Enzyme co-immobilization: Always the biocatalyst designers' choice...or not? *Biotechnol. Adv.* **51**, 107584 (2021).
24. Schoffelen, S. & Van Hest, J. C. M. Multi-enzyme systems: Bringing enzymes together in vitro. *Soft Matter* **8**, 1736–1746 (2012).
25. Hanefeld, U., Gardossi, L. & Magner, E. Understanding enzyme immobilisation. *Chem. Soc. Rev.* **38**, 453–468 (2009).
26. Bommarius, A. S. & Karau, A. Deactivation of Formate Dehydrogenase (FDH) in solution and at gas-liquid interfaces. *Biotechnol. Prog.* **21**, 1663–1672 (2005).
27. Ferrarotti, S. A. *et al.* Immobilization and stabilization of a cyclodextrin glycosyltransferase by covalent attachment on highly activated glyoxyl-agarose supports. *Biotechnol. Prog.* **22**, 1140–1145 (2006).
28. Mateo, C. *et al.* Glyoxyl agarose: A fully inert and hydrophilic support for immobilization and high stabilization of proteins. *Enzyme Microb. Technol.* **39**, 274–280 (2006).
29. Mateo, C., Fernández-Lorente, G., Abian, O., Fernández-Lafuente, R. & Guisán, J. M. Multifunctional epoxy supports: A new tool to improve the covalent immobilization of proteins. The promotion of physical adsorptions of proteins on the supports before their covalent linkage. *Biomacromolecules* **1**, 739–745 (2000).
30. López-Gallego, F. *et al.* Enzyme stabilization by glutaraldehyde crosslinking of adsorbed proteins on aminated supports. *J. Biotechnol.* **119**, 70–75 (2005).
31. Betancor, L. *et al.* Different mechanisms of protein immobilization on glutaraldehyde activated supports: Effect of support activation and immobilization conditions. *Enzyme Microb. Technol.* **39**, 877–882 (2006).
32. Zeballos, N., Diamanti, E., Benítez-Mateos, A. I., Schmidt-Dannert, C. & López-Gallego, F. Solid-Phase Assembly of Multienzyme Systems into Artificial Cellulosomes. *Bioconjug. Chem.* **32**, 1966–1972 (2021).
33. Santiago-arcos, J., Velasco-lozano, S. & López-gallego, F. Multienzyme Coimmobilization on Triheterofunctional Supports. *Biomac.* **24**, 929–94 (2023).

34. Velasco-Lozano, S., da Silva, E. S., Llop, J. & López-Gallego, F. Sustainable and Continuous Synthesis of Enantiopure α -Amino Acids by Using a Versatile Immobilised Multienzyme System. *ChemBioChem* **19**, 395–403 (2018).
35. Breger, J. C. *et al.* Self assembling nanoparticle enzyme clusters provide access to substrate channeling in multienzymatic cascades. *Nat. Commun.* **14**, 1757 (2023).
36. Wang, S. Z. *et al.* Strategies and perspectives of assembling multi-enzyme systems. *Crit. Rev. Biotechnol.* **37**, 1024–1037 (2017).
37. Wang, S., Wang, J., Zhou, X., Guo, Y. & Fang, B. The improvement of stability, activity, and substrate promiscuity of glycerol dehydrogenase substituted by divalent metal ions. *Biotechnol. Bioprocess Eng.* **18**, 796–800 (2013).
38. Perrault, S. D. & Shih, W. M. *Lipid membrane encapsulation of a 3D DNA nano octahedron. Methods in Molecular Biology.* **1500**, 165-184 (2017).
39. Simmel, F. C. DNA-based assembly lines and nanofactories. *Curr. Opin. Biotechnol.* **23**, 516–521 (2012).
40. Nojima, T. *et al.* Nano-Scale Alignment of Proteins on a Flexible DNA Backbone. *PLoS One* **7**, 1–7 (2012).
41. Klein, W. P. *et al.* Enhanced Catalysis from Multienzyme Cascades Assembled on a DNA Origami Triangle. *ACS Nano* **13**, 13677–13689 (2019).
42. Gad, S. & Ayakar, S. Protein scaffolds: A tool for multi-enzyme assembly. *Biotechnol. Reports* **32**, e00670 (2021).
43. Zhang, G., Quin, M. B. & Schmidt-Dannert, C. Self-Assembling Protein Scaffold System for Easy in Vitro Coimmobilization of Biocatalytic Cascade Enzymes. *ACS Catal.* **8**, 5611–5620 (2018).
44. Zhang, G., Johnston, T., Quin, M. B. & Schmidt-Dannert, C. Developing a Protein Scaffolding System for Rapid Enzyme Immobilization and Optimization of Enzyme Functions for Biocatalysis. *ACS Synth. Biol.* **8**, 1867–1876 (2019).
45. Vanderstraeten, J. & Briers, Y. Synthetic protein scaffolds for the colocalisation of co-acting enzymes. *Biotechnol. Adv.* **44**, 107627 (2020).
46. Kummer, M. J. *et al.* Substrate Channeling by a Rationally Designed Fusion Protein in a Biocatalytic Cascade. *JACS Au* **1**, 1187–1197 (2021).
47. Rapali, P. *et al.* Scaffold-mediated gating of cdc42 signalling flux. *Elife* **6**, 1–18 (2017).
48. Bhattacharyya, R. P., Reményi, A., Yeh, B. J. & Lim, W. A. Domains, motifs, and scaffolds: The role of modular interactions in the evolution and wiring of cell signaling circuits. *Annu. Rev. Biochem.* **75**, 655–680 (2006).
49. Gad, S. & Ayakar, S. Protein scaffolds: A tool for multi-enzyme assembly. *Biotechnol. Reports* **32**, e00670 (2021).
50. Chado, G. R., Stoykovich, M. P. & Kaar, J. L. Role of Dimension and Spatial Arrangement on the Activity of Biocatalytic Cascade Reactions on Scaffolds. *ACS Catal.* **6**, 5161–5169 (2016).
51. Idan, O. & Hess, H. Engineering enzymatic cascades on nanoscale scaffolds. *Curr.*

References

- Opin. Biotechnol.* **24**, 606–611 (2013).
52. Fu, J. *et al.* Multi-enzyme complexes on DNA scaffolds capable of substrate channelling with an artificial swinging arm. *Nat. Nanotechnol.* **9**, 531–536 (2014).
53. Dueber, J. E. *et al.* Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* **27**, 753–759 (2009).
54. Zhang, Y. H. P. Substrate channeling and enzyme complexes for biotechnological applications. *Biotechnol. Adv.* **29**, 715–725 (2011).
55. Idan, O. & Hess, H. Origins of activity enhancement in enzyme cascades on scaffolds. *ACS Nano* **7**, 8658–8665 (2013).
56. Zhang, Y. H. P. Substrate channeling and enzyme complexes for biotechnological applications. *Biotechnol. Adv.* **29**, 715–725 (2011).
57. Abdallah, W., Hong, X., Banta, S. & Wheeldon, I. Microenvironmental effects can masquerade as substrate channelling in cascade biocatalysis. *Curr. Opin. Biotechnol.* **73**, 233–239 (2022).
58. Zhang, Y. *et al.* Peptide-mediated immobilization on magnetoferritin for enzyme recycling. *Nanomaterials* **9**, 1–12 (2019).
59. Romero-Fernández, M. & Paradisi, F. Protein immobilization technology for flow biocatalysis. *Curr. Opin. Chem. Biol.* **55**, 1–8 (2020).
60. Bilal, M. & Iqbal, H. M. N. Tailoring Multipurpose Biocatalysts via Protein Engineering Approaches: A Review. *Catal. Letters* **149**, 2204–2217 (2019).
61. Peer, A., Smith, S. P., Bayer, E. A., Lamed, R. & Borovok, I. Noncellulosomal cohesin- and dockerin-like modules in the three domains of life. *FEMS Microbiol. Lett.* **291**, 1–16 (2009).
62. Karpol, A. *et al.* Engineering a reversible, high-affinity system for efficient protein purification based on the cohesin-dockerin interaction. *J. Mol. Recognit.* **22**, 91–98 (2009).
63. Smith, M. R. *et al.* Elucidating structure–performance relationships in whole-cell cooperative enzyme catalysis. *Nat. Catal.* **2**, 809–819 (2019).
64. Liu, F., Banta, S. & Chen, W. Functional assembly of a multi-enzyme methanol oxidation cascade on a surface-displayed trifunctional scaffold for enhanced NADH production. *Chem. Commun.* **49**, 3766–3768 (2013).
65. Tippmann, S. *et al.* Affibody scaffolds improve sesquiterpene production in *saccharomyces cerevisiae*. *ACS Synth. Biol.* **6**, 19–28 (2017).
66. Han, G. H. *et al.* Leucine zipper-mediated targeting of multi-enzyme cascade reactions to inclusion bodies in *Escherichia coli* for enhanced production of 1-butanol. *Metab. Eng.* **40**, 41–49 (2017).
67. Park, S. Y., Eun, H., Lee, M. H. & Lee, S. Y. Metabolic engineering of *Escherichia coli* with electron channelling for the production of natural products. *Nat. Catal.* **5**, 726–737 (2022).
68. Khairil Anuar, I. N. A. *et al.* Spy&Go purification of SpyTag-proteins using pseudo-SpyCatcher to access an oligomerization toolbox. *Nat. Commun.* **10**, 1–13 (2019).

69. Hatlem, D., Trunk, T., Linke, D. & Leo, J. C. Catching a SPY: Using the SpyCatcher-SpyTag and related systems for labeling and localizing bacterial proteins. *Int. J. Mol. Sci.* **20**, (2019).
70. Veggiani, G. *et al.* Programmable polyproteins built using twin peptide superglues. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 1202–1207 (2016).
71. Lawrie, J., Song, X., Niu, W. & Guo, J. A high throughput approach for the generation of orthogonally interacting protein pairs. *Sci. Rep.* **8**, 4–13 (2018).
72. Chen, R. *et al.* Biomolecular scaffolds for enhanced signaling and catalytic efficiency. *Curr. Opin. Biotechnol.* **28**, 59–68 (2014).
73. Andrade, M. A., Perez-Iratxeta, C. & Ponting, C. P. Protein repeats: Structures, functions, and evolution. *J. Struct. Biol.* **134**, 117–131 (2001).
74. Mejias, S. H., Aires, A., Couleaud, P. & Cortajarena, A. L. Designed repeat proteins as building blocks for nanofabrication. *Adv. Exp. Med. Biol.* **940**, 61–81 (2016).
75. Gradišar, H. & Jerala, R. Self-assembled bionanostructures: Proteins following the lead of DNA nanostructures. *J. Nanobiotechnology* **12**, 1–9 (2014).
76. Romera, D., Couleaud, P., Mejias, S. H., Aires, A. & Cortajarena, A. L. Biomolecular templating of functional hybrid nanostructures using repeat protein scaffolds. *Biochem. Soc. Trans.* **43**, 825–831 (2015).
77. D'Andrea, L. D. & Regan, L. TPR proteins: The versatile helix. *Trends Biochem. Sci.* **28**, 655–662 (2003).
78. Binz, H. K., Stumpp, M. T., Forrer, P., Amstutz, P. & Plückthun, A. Designing repeat proteins: Well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins. *J. Mol. Biol.* **332**, 489–503 (2003).
79. Parker, R., Mercedes-Camacho, A. & Grove, T. Z. Consensus design of a NOD receptor leucine rich repeat domain with binding affinity for a muramyl dipeptide, a bacterial cell wall fragment. *Protein Sci.* **23**, 790–800 (2014).
80. Main, E. R. G., Xiong, Y., Cocco, M. J., D'Andrea, L. & Regan, L. Design of stable α -helical arrays from an idealized TPR motif. *Structure* **11**, 497–508 (2003).
81. Cortajarena, A. L. & Regan, L. Ligand binding by TPR domains. *Protein Sci.* **15**, 1193–1198 (2006).
82. Cortajarena, A. L., Liu, T. Y., Hochstrasser, M. & Regan, L. Designed proteins to modulate cellular networks. *ACS Chem. Biol.* **5**, 545–552 (2010).
83. Mejías, S. H. *et al.* Repeat protein scaffolds: Ordering photo- and electroactive molecules in solution and solid state. *Chem. Sci.* **7**, 4842–4847 (2016).
84. López-Andarias, J. *et al.* Toward Bioelectronic Nanomaterials: Photoconductivity in Protein–Porphyrin Hybrids Wrapped around SWCNT. *Adv. Funct. Mater.* **28**, 1–8 (2018).
85. Mejias, S. H. *et al.* Repeat proteins as versatile scaffolds for arrays of redox-active FeS clusters. *Chem. Commun.* **55**, 3319–3322 (2019).
86. Mejias, S. H. *et al.* Engineering conductive protein films through nanoscale self-assembly and gold nanoparticles doping. *Nanoscale* **13**, 6772–6779 (2021).

References

87. Mejías, S. H., Sot, B., Guantes, R. & Cortajarena, A. L. Controlled nanometric fibers of self-assembled designed protein scaffolds. *Nanoscale* **6**, 10982–10988 (2014).
88. Phillips, J. J., Millership, C. & Main, E. R. G. Fibrous nanostructures from the self-assembly of designed repeat protein modules. *Angew. Chemie - Int. Ed.* **51**, 13132–13135 (2012).
89. Salgado, E. N., Lewis, R. A., Mossin, S., Rheingold, A. L. & Tezcan, F. A. Control of protein oligomerization symmetry by metal coordination: C₂ and C₃ symmetrical assemblies through Cu II and Ni II coordination. *Inorg. Chem.* **48**, 2726–2728 (2009).
90. Salgado, E. N., Radford, R. J. & Tezcan, F. A. Metal-directed protein self-assembly. *Acc. Chem. Res.* **43**, 661–672 (2010).
91. Bailey, J. B., Subramanian, R. H., Churchfield, L. A. & Tezcan, F. A. Metal-Directed Design of Supramolecular Protein Assemblies. *Methods in Enzymology.* **580**, 223–250 (2016).
92. Brodin, J. D. *et al.* Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. *Nat. Chem.* **4**, 375–382 (2012).
93. Speltz, E. B., Nathan, A. & Regan, L. Design of Protein-Peptide Interaction Modules for Assembling Supramolecular Structures in Vivo and in Vitro. *ACS Chem. Biol.* **10**, 2108–2115 (2015).
94. Kang, W. *et al.* Modular enzyme assembly for enhanced cascade biocatalysis and metabolic flux. *Nat. Commun.* **10**, (2019).
95. Guo, Q. *et al.* Structural and Kinetic Studies of Formate Dehydrogenase from *Candida boidinii*. *Biochemistry* **55**, 2760–2771 (2016).
96. Hayward, S. & Kitao, A. Molecular dynamics simulations of NAD⁺-induced domain closure in horse liver alcohol dehydrogenase. *Biophys. J.* **91**, 1823–1831 (2006).
97. Velasco-Lozano, S., Jackson, E., Ripoll, M., López-Gallego, F. & Betancor, L. Stabilization of ω-transaminase from *Pseudomonas fluorescens* by immobilization techniques. *Int. J. Biol. Macromol.* **164**, 4318–4328 (2020).
98. Lin, Y., Leng, R. P. & Benchimol, S. Identification of p53-regulated genes by the method of differential display. *Methods Mol Biol.* **317**, 193–206 (2006).
99. Brinker, A. *et al.* Ligand discrimination by TPR domains. Relevance and selectivity of EEVD-recognition in Hsp70-Hop-Hsp90 complexes. *J. Biol. Chem.* **277**, 19265–19275 (2002).
100. Crowe, J., Masone, B. S. & Ribbe, J. One-step purification of recombinant proteins with the 6xHis tag and Ni-NTA resin. *Methods Mol. Biol.* **58**, 491–510 (1996).
101. Cohen, S. L. & Chait, B. T. Mass spectrometry of whole proteins eluted from sodium dodecyl sulfate- polyacrylamide gel electrophoresis gels. *Anal. Biochem.* **247**, 257–267 (1997).
102. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* **1**, 2876–2890 (2007).
103. Miles, A. J. & Wallace, B. A. Circular Dichroism Spectroscopy for Protein Characterization: Biopharmaceutical Applications. *Biophys. Charact. Proteins Dev.*

- Biopharm.* 109-137 (2015).
104. Samarina, N. *et al.* Recruitment of phospholipase C γ 1 to the nonstructural membrane protein pK15 of Kaposi Sarcoma-associated herpesvirus promotes its Src-dependent phosphorylation. *PLoS Pathog.* **17**, 1–30 (2021).
 105. Aires, A. *et al.* A Simple Approach to Design Proteins for the Sustainable Synthesis of Metal Nanoclusters. *Angew. Chemie - Int. Ed.* **58**, 6214–6219 (2019).
 106. Boeckx, J., Hertog, M., Geeraerd, A. & Nicolai, B. Kinetic modelling: An integrated approach to analyze enzyme activity assays. *Plant Methods* **13**, 1–12 (2017).
 107. Srinivasan, B. A guide to the Michaelis–Menten equation: steady state and beyond. *FEBS J.* **289**, 6086-6089 (2022).
 108. Walsh, R. Are improper kinetic models hampering drug development? *PeerJ* **2014**, (2014).
 109. Bhushan, R. & Brückner, H. Marfey's reagent for chiral amino acid analysis: A review. *Amino Acids* **27**, 231–247 (2004).
 110. Grove, T. Z., Hands, M. & Regan, L. Creating novel proteins by combining design and selection. *Protein Eng. Des. Sel.* **23**, 449–455 (2010).
 111. Wu, X. *et al.* Construction and characterization of novel bifunctional fusion proteins composed of alcohol dehydrogenase and NADH oxidase with efficient oxidized cofactor regeneration. *Biotechnol. Appl. Biochem.* **69**, 1535-1544 (2021).
 112. Mantoanelli, J. O. F., Gonçalves, L. M. & Pereira, E. A. Dansyl Chloride as a Derivatizing Agent for the Analysis of Biogenic Amines by CZE-UV. *Chromatographia* **83**, 767–778 (2020).
 113. Rai, V. & Dey, N. The Basics of Confocal Microscopy. *Laser Scanning, Theory Appl.* (2011).
 114. Chen, K. & Arnold, F. H. Engineering new catalytic activities in enzymes. *Nat. Catal.* **3**, 203–213 (2020).
 115. Aalbers, F. S. & Fraaije, M. W. Enzyme Fusions in Biocatalysis: Coupling Reactions by Pairing Enzymes. *ChemBioChem* **20**, 20–28 (2019).
 116. Yu, K., Liu, C., Kim, B. G. & Lee, D. Y. Synthetic fusion protein design and applications. *Biotechnol. Adv.* **33**, 155–164 (2015).
 117. Monterrey, D. T., Ayuso-Fernández, I., Oroz-Guinea, I. & García-Junceda, E. Design and biocatalytic applications of genetically fused multifunctional enzymes. *Biotechnol. Adv.* **60**, 108016 (2022).
 118. Mechaly, A. *et al.* Cohesin-dockerin recognition in cellulosome assembly: Experiment versus hypothesis. *Proteins Struct. Funct. Genet.* **39**, 170–177 (2000).
 119. Ledesma-fernandez, A., Velasco-lozano, S., Santiago-arco, J., López-gallego, F. & Cortajarena, A. L. Engineered repeat proteins as scaffolds to assemble multi-enzyme systems for efficient cell-free biosynthesis *Nat Commun* **14**, 2587 (2023).
 120. Cortajarena, A. L., Mochrie, S. G. J. & Regan, L. Modulating repeat protein stability: The effect of individual helix stability on the collective behavior of the ensemble. *Protein Sci.* **20**, 1042–1047 (2011).

References

121. Cortajarena, A. L. & Regan, L. Calorimetric study of a series of designed repeat proteins: Modular structure and modular folding. *Protein Sci.* **20**, 336–340 (2011).
122. Grove, T. Z., Regan, L. & Cortajarena, A. L. Nanostructured functional films from engineered repeat proteins. *J. R. Soc. Interface* **10**, (2013).
123. Uribe, K. B. *et al.* Engineered Repeat Protein Hybrids: The New Horizon for Biologic Medicines and Diagnostic Tools. *Acc. Chem. Res.* **54**, 4166–4177 (2021).
124. Cortajarena, A. L., Yi, F. & Regan, L. Designed TPR modules as novel anticancer agents. *ACS Chem. Biol.* **3**, 161–166 (2008).
125. Aires, A., Möller, M. & Cortajarena, A. L. Protein Design for the Synthesis and Stabilization of Highly Fluorescent Quantum Dots. *Chem. Mater.* **32**, 5729–5738 (2020).
126. Kajander, T., Cortajarena, A. L., Mochrie, S. & Regan, L. Structure and stability of designed TPR protein superhelices: Unusual crystal packing and implications for natural TPR proteins. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **63**, 800–811 (2007).
127. Cortajarena, A. L., Aires, A., Sousaraei, A., Möller, M. & Cabanillas-Gonzalez, J. Boosting the photoluminescent properties of protein-stabilized gold nanoclusters through protein engineering. *Nano Lett.* **21**, 9347–9353 (2021).
128. Zheng, H. *et al.* CheckMyMetal: A macromolecular metal-binding validation tool. *Acta Crystallogr. Sect. D Struct. Biol.* **73**, 223–233 (2017).
129. Bauler, P., Huber, G., Leyh, T. & McCammon, J. A. Channeling by proximity: The catalytic advantages of active site colocalization using brownian dynamics. *J. Phys. Chem. Lett.* **1**, 1332–1335 (2010).
130. Wheeldon, I. *et al.* cascade reactions. *Nat. Publ. Gr.* **8**, 299–309 (2016).
131. Castellana, M. *et al.* Enzyme clustering accelerates processing of intermediates through metabolic channeling. *Nat. Biotechnol.* **32**, 1011–1018 (2014).
132. Aires, A., Lopez-Martinez, E. & Cortajarena, A. L. Sensors based on metal nanoclusters stabilized on designed proteins. *Biosensors* **8**, (2018).
133. Rosenthal, K., Bornscheuer, U. T. & Lütz, S. Cascades of Evolved Enzymes for the Synthesis of Complex Molecules. *Angew. Chemie Int. Ed.* **202208358**, 4–6 (2022).
134. You, C. & Zhang, Y. H. P. Self-assembly of synthetic metabolons through synthetic protein scaffolds: One-step purification, co-immobilization, and substrate channeling. *ACS Synth. Biol.* **2**, 102–110 (2013).
135. Goldhahn, C., Burgert, I. & Chanana, M. Nanoparticle-Mediated Enzyme Immobilization on Cellulose Fibers: Reusable Biocatalytic Systems for Cascade Reactions. *Adv. Mater. Interfaces* **6**, 1–9 (2019).
136. Lin, J. L., Zhu, J. & Wheeldon, I. Synthetic Protein Scaffolds for Biosynthetic Pathway Colocalization on Lipid Droplet Membranes. *ACS Synth. Biol.* **6**, 1534–1544 (2017).
137. Adams, J. J., Pal, G., Jia, Z. & Smith, S. P. Mechanism of bacterial cell-surface attachment revealed by the structure of cellulosomal type II cohesin-dockerin complex. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 305–310 (2006).

138. Eun, C., Kekenés-Huskey, P. M., Metzger, V. T. & McCammon, J. A. A model study of sequential enzyme reactions and electrostatic channeling. *J. Chem. Phys.* **140**, (2014).
139. Lin, J. & Wheeldon, I. Kinetic Enhancements in DNA – Enzyme Nanostructures Mimic the Sabatier Principle. *ACS Catal.* **3**, 560-564 (2013).
140. Zhang, Y., Tsitkov, S. & Hess, H. peroxidase cascade. *Nat. Commun.* **7**, 1–9 (2016).
141. Sen, A. & Kohen, A. Enzymatic tunneling and kinetic isotope effects: Chemistry at the crossroads. *J. Phys. Org. Chem.* **23**, 613–619 (2010).
142. Lay, J. P. & Hammes-schi, S. Hydrogen Tunneling in Enzymes and Biomimetic Models. *Chem. Rev.* **114**, 3466–3494 (2014).
143. Doménech, A. *et al.* A thermodynamic , electrochemical and molecular dynamics study on NAD and NADP recognition by 1 , 4 , 7 , 10 , 13 , 16 , 19-heptaaza-cycloheptacosane ([21] aneN 7) †. **4**, 23–32 (1999).
144. Morlock, L. K., Böttcher, D. & Bornscheuer, U. T. Simultaneous detection of NADPH consumption and H₂O₂ production using the Ampliflu™ Red assay for screening of P450 activities and uncoupling. *Appl. Microbiol. Biotechnol.* **102**, 985–994 (2018).
145. Schindelin, J. *et al.* Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
146. Bolte, S. & Cordelières, F. P. A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* **224**, 213–232 (2006).
147. Adler, J. & Parmryd, I. Quantifying colocalization by correlation: The pearson correlation coefficient is superior to the Mander's overlap coefficient. *Cytom. Part A* **77**, 733–742 (2010).
148. Simon, D. S., Sergienko, A. V, Mary, S. & Ave, C. The correlation confocal microscope. **18**, 9765–9779 (2010).
149. Velasco-Lozano, S., Santiago-Arcos, J., Mayoral, J. A. & López-Gallego, F. Co-immobilization and Colocalization of Multi-Enzyme Systems for the Cell-Free Biosynthesis of Aminoalcohols. *ChemCatChem* **12**, 3030–3041 (2020).
150. Guo, F. & Berglund, P. Transaminase biocatalysis: Optimization and application. *Green Chem.* **19**, 333–360 (2017).
151. Meier, R. J., Gundersen, M. T., Woodley, J. M. & Schürmann, M. A Practical and Fast Method to Predict the Thermodynamic Preference of ω -Transaminase-Based Transformations. *ChemCatChem* **7**, 2594–2597 (2015).
152. Gundersen, M. T., Abu, R., Schürmann, M. & Woodley, J. M. Amine donor and acceptor influence on the thermodynamics of ω -transaminase reactions. *Tetrahedron Asymmetry* **26**, 567–570 (2015).
153. Schätzle, S., Höhne, M., Redestad, E., Robins, K. & Bornscheuer, U. T. Rapid and sensitive kinetic assay for characterization of ω -transaminases. *Anal. Chem.* **81**, 8244–8248 (2009).
154. Mutti, F. G., Fuchs, C. S., Pressnitz, D., Sattler, J. H. & Kroutil, W. Stereoselectivity of four (R)-selective transaminases for the asymmetric amination of ketones. *Adv. Synth. Catal.* **353**, 3227–3233 (2011).

References

155. Li, Y. & Cirino, P. C. Recent advances in engineering proteins for biocatalysis. *Biotechnol. Bioeng.* **111**, 1273–1287 (2014).
156. Sánchez-deAlcázar, D., Velasco-Lozano, S., Zeballos, N., López-Gallego, F. & Cortajarena, A. L. Biocatalytic Protein-Based Materials for Integration into Energy Devices. *ChemBioChem* **20**, 1977–1985 (2019).
157. Rodríguez-Abetxuko, A., Sánchez-deAlcázar, D., Cortajarena, A. L. & Beloqui, A. A Versatile Approach for the Assembly of Highly Tunable Biocatalytic Thin Films. *Adv. Mater. Interfaces* **6**, 1–9 (2019).

Abbreviations

Abbreviations

DNA	Deoxyribonucleic acid
TPR	Tetratricopeptide repeat protein
CTPR	Consensus tetratricopeptide repeat protein
SCAB	Scaffolding bio-brick
TRAP	Tetratricopeptide repeat affinity protein
FDH	Formate dehydrogenase enzyme
LAlaDH	L-Alanine dehydrogenase enzyme
PDB	Protein databank file
NADH/NAD ⁺	Nicotinamide adenine dinucleotide
SDS-PAGE gel	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
ω TA	ω -transaminase enzyme
K _D	Dissociation constant
EutM	Ethanolamine bacterial microcompartment protein
ANK	Ankyrin repeat proteins
LLR	Leucine-rich repeat proteins
C, Cys	Cysteine amino acid
H, His	Histidine amino acid
WT	Wild-type protein
CTPR3 WT	CTPR with three repeats
PCR	Polymerase chain reaction
CTPR1 WT	CTPR with 1 repeat
QC	Quick change Site-Directed mutagenesis
HOP protein	Heat-shock protein
LB	Luria-Bertain broth
IPTG	Isopropyl thiol- β -galactosidase
DTT	Dithiothreitol
Ni-NTA	Nickel nitriloacetic affinity chromatography
TEV protease	Tobacco Etch virus protease
EDTA	Ethylenediaminetetraacetic acid

Abbreviations

SEC	Size exclusion chromatography
FPLC	Fast protein liquid chromatography
MALDI-TOF	Matrix-Assisted Laser Desorption/Ionization-Time-Of-Flight
PBS buffer	Phosphate-buffered saline
TRIS-HCl	Tris(hydroxymethyl) aminomethane hydrochloride
TFA	Trifluoroacetic acid
CD	Circular dichroism
T_M	Melting temperature
DMSO	Dimethyl sulfoxide
PD-10	Disposable desalting column
Rh	Hydrodynamic radius
ICP-MS	Inductively coupled plasma mass spectrometry
UV-Vis	Ultraviolet visible spectroscopy
FEA	Methylbenzylamine
PLP	Pyridoxal 5'-phosphate monohydrate
HEPES buffer	4-(2-Hydroxyethyl) piperazine-1-Ethanesulfonic Acid
K_M	Michaelis-Menten constant
V_{max}	Maximum rate
K_{cat}	Turnover number
K_{cat}/K_M	Catalytic efficiency
K_i	Inhibition constant
HPLC	High Performance Liquid Chromatography
NOX	NADH oxidase enzyme
FAD ⁺	Flavin adenine dinucleotide
HRP	Horseradish peroxidase enzyme
AR	Amplex red
RSF	Resorufin
ϵ	Extinction coefficient
UV-Vis	Ultraviolet visible spectroscopy

Abbreviations

UPLC-MS	Ultra-Performance Liquid Chromatography-Mass Spectrometry
ESI-TOF	Electrospray ionization-Time-of-flight Mass Spectrometry
AF647	Alexa fluor 647 fluorescent dye
AF488	Alexa fluor 488 fluorescent dye
ORF	Open reading frame
MW	Molecular weight
CY	Chromatographic yield
AG-Co ²⁺	Cobalt agarose resin
CLSM	Confocal laser scanning microscopy
PCC	Pearson correlation coefficient
nMDP	Normalized mean deviation product
MOC	Manders overlap coefficient
MA/MB	Manders coefficients
CTPR20	CTPR with 20 repeats
CTPR8	CTPR with 8 repeats
Cu, Ni, Co	Copper, nickel, and cobalt metals
CMM	Check my metal server

Acknowledgments

I look back and today I still can't believe that that girl who started at the university of chemistry full of fears has become a doctor. I must admit that it was not something premeditated that I had in mind from the beginning like other colleagues. But as the years went by, things happened, and destiny took me here. Everything has a reason, don't they say?

First, I would like to express my sincere thanks to my advisors, Dr. Aitziber López Cortajarena and Dr. Fernando López Gallego, for giving me the opportunity to do my thesis within the European HOMBIOCAT project under their supervision on a topic which I have enjoyed very much. I would also like to thank them for their continuous support in my Ph.D. study and research, for their patience, motivation, enthusiasm, and immense knowledge. Your guidance helped me throughout my research and writing of this thesis.

Furthermore, I wish to show my gratitude to the CIC biomaGUNE centre and all the staff members for their support and excellent work, from administration/maintenance to the different platforms (especially to Irantzu and Javi for the training and help you have given me to learn about different techniques that have been fundamental to move forward with my thesis).

All my gratitude to each one of my colleagues in the laboratory of the biomolecular nanotechnology group. For the good and not so good moments that we have lived together during these 4 years: the meals, the coffees, even the group meetings, ha, ha. Thank you for making this journey much easier and more pleasant. To the golden triplet, Gabriela, Liher and Rocío. Gabriela, be tenacious and you will win. Thanks for your wise advice and bad jokes to make me laugh. Liher, who would have thought that I would end up working hand in hand with a Bilbaino, what is life. Eta azkenik, nire hitzik onenak gordetzen ditut nire euskaldun-andaluziarrarentzat, Rocío. Ez nuen inoiz pentsatu doktoradutzan benetako lagun bat egin nezakeenik. Beste inork baino hobeto ulertu nauzu, eta ezinbesteko pieza izan zara bide honetan zehar. Zure aholkuak, zure pazientzia eta baita ere zure adiskidetasuna ezinbestekoak izan dira honaino iristeko. Badakit beti hor egongo zarela eta ez nauzula ahaztuko. Gora katalisiaren boterea!!!! I would also like to highlight the people of the HetBioCat laboratory, Javi and Susana, thanks to whom I have learned a lot about the world of biocatalysis, and who are an indispensable part of the results obtained in my thesis.

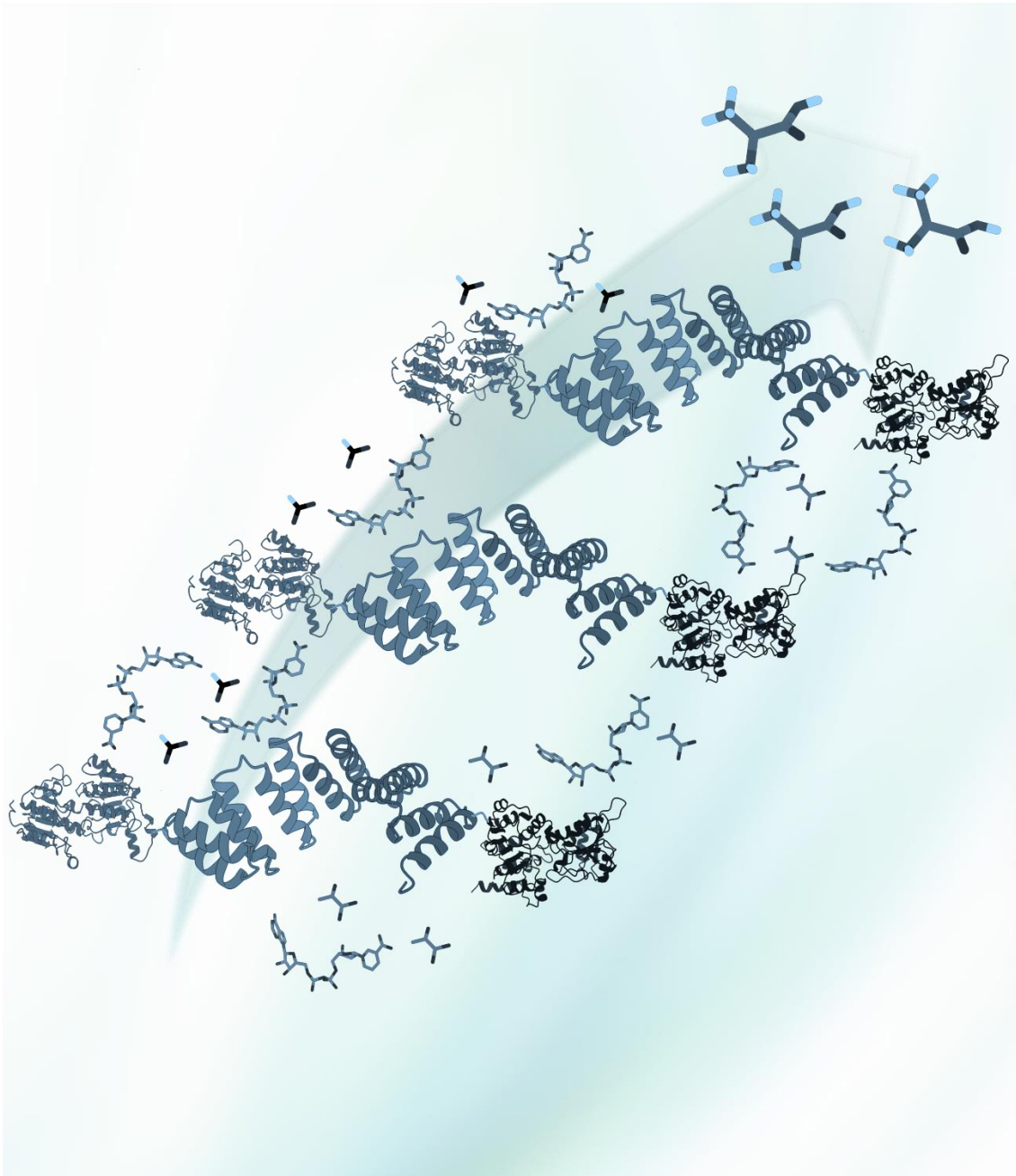
I don't want to forget my other little scientific family; "parte alante and triplete". Thank you, girls, because thanks to you in the five years we spent at the university together not only made it a really enriching experience, but also gave me the bug to continue in the world of science. I would like to thank Julia, who has become my battle companion and indispensable friend. You better than anyone know what we have been through and look how we are now, who would have thought it eh.

Por último, pero no menos importante, quiero agradecer a mi familia y a mis amigas: mis padres Rosi e Iñaki y a mi hermano, Iker. Sois lo más importante para mí y sin vosotros nada de esto habría sido posible. A mis amigas Hasbulla team: Leila, Laura, Saray y Janira. Que deciros a vosotras que no sepáis. Simplemente gracias y más gracias por hacerme los días más amenos: las charlas, las fiestas, los viajes, los cafés... por entenderme siempre en los buenos y malos momentos y ser mis compañeras de

Acknowledgments

vida. Finally, I want to say that even though the road has been long and hard; that there are good days and bad days, it has been an experience that I would recommend to anyone who is passionate about science, especially to those girls, that we are more and more women scientists, and we can achieve great things too. I reaffirm that this stage has helped me to mature and give importance to the important things.

To finish by saying that *esto recién empieza...* and there is still the funniest part of the story.



eman ta zabal zazu



Universidad
el País Vasco

Euskal Herriko
Unibertsitatea

CICbiomaGUNE

MEMBER OF BASQUE RESEARCH
& TECHNOLOGY ALLIANCE

