RSC Advances

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *RSC Adv.*, 2016, **6**, 208

Received 13th November 2015
Accepted 11th December 2015

DOI: 10.1039/c5ra23949d

www.rsc.org/advances

Protein self-assembly induces promiscuous nucleophilic biocatalysis in Morita–Baylis–Hillman (MBH) reaction†

Pralhad N. Joshi, Landa Purushottam, Nirmal K. Das, Saptarshi Mukherjee and Vishal Rai*

Self-assembled states of proteins render efficient promiscuous nucleophilic biocatalysis in MBH reaction in a green process. The His and Arg based catalophores in proteins operate in aqueous buffer at neutral pH and ambient temperature. Steady-state fluorimetric approaches reveal that lower order aggregates play a seminal role in the biocatalytic process.

The chemical industry has seen a remarkable increase in the use of biocatalysis in recent years.[1] Structurally defined spatial distribution of residues in proteins has offered efficient catalysis for a variety of mechanistically unique synthetic transformations.[2–4] Over a period, the perspective of high specificity in enzyme catalysis has changed and led to the realization that a large population of enzymes is capable of catalyzing multiple reactions.[5] However, non-biological transformations have posed a stiff challenge to biocatalysis. Promiscuous enzymes or protein-based catalysts are key aspirants in this case and have drawn attention from diverse streams of Science.[6] One of the most important synthetic transformations, carbon–carbon bond formation, is realized in natural systems[7] but do not extend to non-biological chemical transformations. For example, MBH reaction has been a notoriously difficult target owing to a sequence of steps that involves mechanism switching based on the substrates and solvents. MBH reaction involves nucleophilic catalysis in a Michael addition rendering enolate that entraps an electrophile to form C–C bond in an aldol condensation and results in a functionally rich molecule. There are no examples of antibodies that can catalyze this reaction. A recent attempt in this direction through directed evolution of engineered enone-binding proteins[8] and another previous attempt[9] met limited success. Protein backbone has several nucleophilic residues where histidine (His) serves as a potential candidate for nucleophilic catalysis. However, the earlier efforts (Fig. 1a) with His or His tag bearing proteins and our initial investigations (Table S1, ESI†) suggested that His alone is not suitable for catalyzing the MBH reaction.

Here, we report promiscuous catalysis by lysozyme C **5c** and myoglobin **5e** hinged on catalophore formed by His and Arginine (Arg) pair to catalyze the MBH reaction with remarkably high efficiency. A set of control experiments (Table S1, ESI†) suggested that Arg can potentially mediate proton shuttling process in the transition state[10] (**4**) and assist His for the necessary activation (Fig. 1b). The fluorimetric investigations indicate that the lower order assembly of catalophore consisting protein serves as an active catalyst. These results invite reconsideration of the catalytic species in the field of biocatalysis and pave the path for its potential integration with protein aggregation.

Initially, we selected a variety of proteins (**5a–5e**, Table 1) that offer His and Arg pair confined variably in the conformational space. Cyclohexenone **2** and *p*-nitrobenzaldehyde **1a** were selected as Michael acceptor and electrophile, respectively. When RNase A **5a** was employed as a catalyst, 35% conversion to MBH adduct was observed (entry 1, Table 1). His and Arg residues in RNase A (PDB ID : 2AAS) are separated by ~11.3 Å in a constrained microenvironment, with a lysine (Lys) side chain

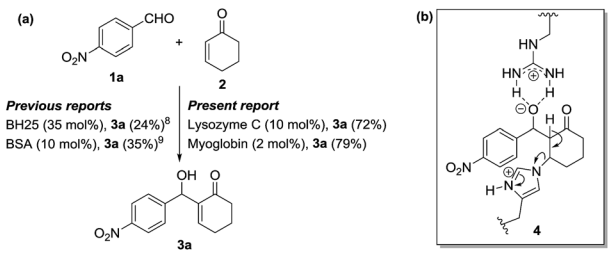
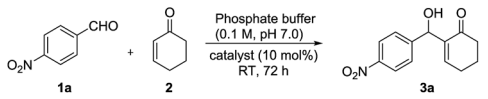


Fig. 1 (a) Biocatalysis in MBH reaction. (b) Hypothetical activation of nucleophile and electrophile in a Michael addition–aldol condensation sequence (MBH) draws comparison with the catalytic activity of lipases.[11]

Department of Chemistry, Indian Institute of Science Education and Research (IISER) Bhopal, Bhopal Bypass Road, Bhauri, Bhopal, MP 462 066, India. E-mail: vrai@iiserb.ac.in

† Electronic supplementary information (ESI) available: Methods, procedures, additional results and discussion, characterization data. See DOI: 10.1039/c5ra23949d

Table 1 Identification of protein that renders His and Arg pair appropriate for MBH reaction



Entry	Catalyst[a]	His and Arg pair	% conversion (3a)[b]
1	RNase A **5a**	1 (H119, R10)	35
2	α-Chymotrypsinogen **5b**	1 (H40, R145)	53
3	Lysozyme C **5c**	1 (H15, R14)	72
4	α-Chymotrypsin **5d**	2 (H91, R93; H65, R217)	37 ± 10
5	Myoglobin **5e**	2 (H24, R118; H113, R31)	65
6	Trypsin **5f**	0	16 ± 8
7	Ubiquitin **5g**	0	22

[a] Protein concentration (2.5 mM). [b] % conversion by ^{1}H NMR. Relative stoichiometry of **1a** and **2** is 1 : 3 after optimization. **1a** and **2** were dissolved in DMSO (buffer : DMSO, 9 : 1) for all the reported reactions in this paper (see ESI for details). For catalyst loading, see Tables S2 and S3 in ESI.

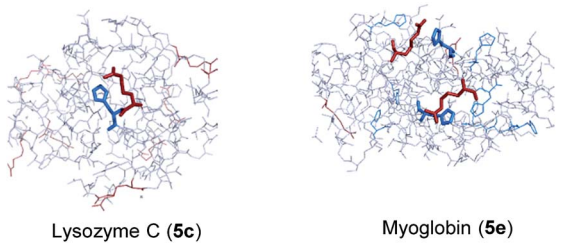


Fig. 2 His (blue) and Arg (red) pair(s) in lysozyme C **5c** and myoglobin **5e**. (See Fig. S1† for **5a**, **5b**, **5d** and **5f–g**).

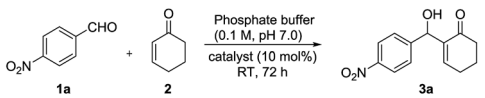
placed between them. α-Chymotrypsinogen **5b** (PDB ID : 1EX3) aligns these residues at ~8.3 Å with no interfering residue and resulted in 18% improvement in the conversion to MBH adduct (entry 2). Lysozyme C **5c** (PDB ID : 2ZYP, Fig. 2), where the two residues are separated only by ~4.2 Å resulted in remarkable improvement with 72% conversion (entry 3, Table 1). The inter-residue distance is only of qualitative interest as solvent accessibility or dynamics of the domain is not taken into account. Although α-chymotrypsin **5d** (PDB ID : 4CHA) and myoglobin **5e** (PDB ID : 1MBO) have two sets of His and Arg pairs, yet in the former, the residues were reasonably separated as compared to the latter. Perhaps, this explains why myoglobin **5e** (Fig. 2) turned out to be a better catalyst than α-chymotrypsin **5d** (entries 4 and 5, Table 1). Besides, we realized that both α-chymotrypsin **5d** and trypsin **5f** result in capricious conversions in multiple attempts due to self-digestion of proteins.

In a control experiment, we selected proteins in which none of the His and Arg residues are in close vicinity (entries 6 and 7). In both the cases, poor reactivity was observed. In yet another control experiment, we blocked the nucleophilic site of His15 in lysozyme C **5c** by a chemoselective transformation with 2,4′-dibromoacetophenone.[12] The modified **5c** resulted in conversions (20%) comparable to **5f** and **5g**, thereby substantiating the role of His in catalysis. It was evident from these results that both the presence and inter-residual distance/orientation between His and Arg are the driving parameters in controlling the efficiency of the catalytic process.

The concentration of the reaction mixture was optimized with respect to the protein (Table 2). The optimal concentration was found to be 2.5 mM for lysozyme C **5c** and 1 mM for myoglobin **5e** (entries 5 and 4). We observed striking non-linear change in efficiency of catalyst from 1 to 2.5 mM for lysozyme C **5c** and 500 μM to 1 mM in case of myoglobin **5e**. The results indicated that protein assembly rendered a constitution of the active catalyst.[13] The aggregation phenomenon of proteins and peptides has been investigated with techniques such as ultracentrifugation,[14] dialysis,[15] Gouy interferometry,[16] neutron scattering,[17] NMR dispersion,[18] NMR diffusion[19] and fluorimetry.[20] A clear mechanistic understanding in this area is still elusive, but there is a consensus about aggregation with respect to the concentration. It outlines that 1.0 mM lysozyme C **5c** solution can be considered unsaturated primarily having monomers whereas, 2.5 mM solution can be regarded as saturated having lysozyme C **5c** assemblies.

We believe that lysozyme C **5c** assemblage leads to the observed catalytic activity and that too at a specific concentration of **5c**, as observed herein. In order to have a better mechanistic understanding of the actual process of aggregation the protein undergoes, we have carried out thioflavin T (ThT) assays at various concentrations of lysozyme C **5c** (0.5–5 mM). The ThT fluorescence spectra in the presence of varying concentrations of myoglobin **5e** were too structured having multiple well-defined peaks. Hence, to avoid the over-interpretation of data owing to the aggregation process, we did not carry out the fluorescence studies with myoglobin further. The fluorescent probe, ThT is unique in the sense that it has been widely used to unravel the nuances associated with the process of self-assembly and/or aggregation of proteins and peptides.[20] The dramatic increment of fluorescence intensity of ThT upon

Table 2 Effect of protein concentration on MBH reaction



		% conversion (3a)[b,c]	
Entry	Concentration[a]	Lysozyme C (**5c**)	Myoglobin (**5e**)
1	50 μM	0	0
2	100 μM	0	0
3	500 μM	8	9
4	1 mM	10	79
5	2.5 mM	72	65
6	5 mM	27	55
7	10 mM	25	30

[a] Concentration with respect to the protein. [b] % conversion determined by ^{1}H NMR. [c] No background reaction was observed in catalyst free condition at all the concentrations.

binding to the β-rich deposits makes it an excellent tool for deciphering self-aggregation phenomenon. It is believed that ThT not only binds to a variety of substrates but also the micellar form (its critical micelle concentration is ~4 μM in water) interacts with the fibrils.[20a] Lysozyme C **5c** in its native state is composed of ~32% α-helix and ~7% β-sheet[21] and due to aggregation, the content of the latter increases dramatically which gets reflected by the rise in fluorescence intensity of ThT (Fig. S2, ESI†).

Fig. 3a represents the variation of fluorescence intensities against time for different concentrations of lysozyme C **5c** used in the present investigation. As seen from the figure, the fluorescence intensity of ThT at any instant rises almost linearly with increasing concentration of the protein that is a clear signature of the concentration-driven oligomerization that the protein undergoes. The oligomeric states are largely stable over a period of 72 hours. It must be stated here that the emission maxima of ThT (Fig. S2, ESI†) is not a function of lysozyme C **5c** concentration, thereby signifying that the microenvironment of the fluorophore remains almost unaltered due to aggregation. The process of aggregation goes beyond the catalytic maxima of 2.5 mM lysozyme C **5c** and the continued rise in fluorescence intensities of ThT beyond that point is a signature of the increased β-sheet content as a result of aggregation. Lysozyme C **5c**, an intrinsically fluorescent protein has six tryptophan (Trp) residues. Out of these, only two Trp62 and Trp108 are dominant and located in the substrate binding site.[22] We also monitored the emission characteristics of Trp as a function of increasing concentration of lysozyme C **5c** (Fig. S3, ESI†) and observed (spectra were recorded till 72 hours of incubation) that the fluorescence intensities of Trp decreases upon increasing the protein concentration (Fig. S4, ESI†). Interestingly, here also, we did not observe any appreciable change in the emission maxima of Trp indicating that even upon aggregation, the fluorophore does not either undergo major internalization or gets exposed to the hydrophilic environment. The gradual fall in the fluorescence intensities of Trp with a rise in lysozyme C **5c** concentration can be attributed to the self-quenching brought in by the additional amino acid residues that are in close proximity of the fluorophore. To have a molecular level interpretation of our steady-state spectroscopic data, we carried out molecular docking studies (Fig. 3b).[23] ThT binds at a distance of 21.9 Å and 24.8 Å from His and Arg residues of catalophore, respectively and the estimated free energy of binding is −7.91 kcal mol^{-1}. Although ThT is located substantially away from the catalytic site, nonetheless, it is capable of reporting the increased concentration-induced aggregation as evidenced from the spectroscopic data. The microenvironment of both ThT and Trp (both being in close vicinity to each other) does not change due to aggregation. Thus, a linear assembly along the *X*–*Y* plane encompassing the catalophore can be envisaged. The increase in lysozyme C **5c** concentration induces higher order aggregation, however, the catalytic site gets deactivated likely due to excessive crowding and hence the catalytic efficiency goes through a maxima (at ~2.5 mM lysozyme C **5c**, entry 5, Table 2). The diameter of **5c** assembly at 2.5 mM was determined to be 10.7 nm (Fig. S6†) by dynamic light scattering (DLS) experiments. Electron microscopy (FE-SEM) was not found suitable to probe these systems due to enhancement in order of aggregation during sample preparation (Fig. S5†).

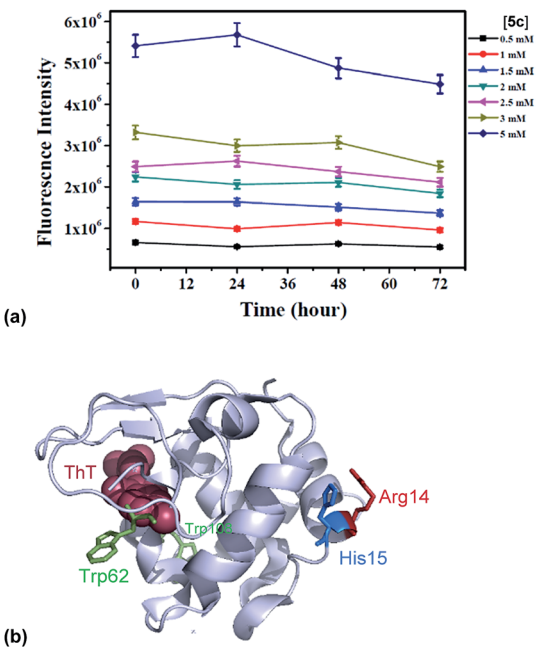
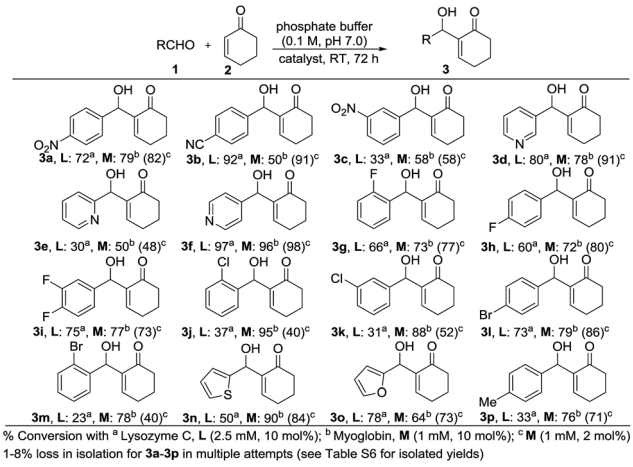


Fig. 3 (a) ThT fluorescence assay with increasing concentration of lysozyme C **5c** as a function of time. The colour black, red, blue, cyan, pink, yellow, and navy blue represents the concentration of lysozyme C 0.5, 1, 1.5, 2, 2.5, 3, 5 mM, respectively. (b) Energy minimized molecular docked structure of lysozyme C with ThT.

The data suggest that the lower order assemblies play a crucial role in catalysis. This is likely enabled by the forged hydrophobic cavities in these protein constructs (lysozyme C **5c** at 2.5 mM) that allow increased effective concentration of reagents. A sharp decrease in efficiency with higher order assembly of **5c** (5 mM) reaffirms that the domain with catalophore is involved in aggregation as complemented by the fluorimetric analysis. A subtle balance of protein concentration is apparently crucial for the success of these catalytic transformations.

With the concentration of protein as a fixed parameter, we varied the amount of catalyst and found 10 mol% of lysozyme C **5c** and 2–10 mol% of myoglobin **5e** were efficient (Table S2, ESI†). The optimized conditions were then examined for their application to a variety of substrates (Scheme 1). It was intriguing to see that lysozyme C **5c** and myoglobin **5e** were able to catalyze MBH reaction with a broad range of electronically diverse aldehydes. The relative catalytic efficiency of lysozyme C **5c** and myoglobin **5e** is dependent on the choice of aldehydes. The latter was superior in multiple cases (**3a**, **3c–e**, **3g–h**, **3j–n** and **3p**) whereas at other instances, both **5c** and **5e** were equally efficient. Interestingly, 2 mol% of myoglobin was more efficient than 10 mol% in two cases (**3b**, **3d**) indicating towards the substrate dependent behaviour of protein assemblies.

It was evident that the reactivity profile offered by these two proteins was unique with respect to other nucleophilic catalysts used in MBH reaction. To substantiate this point, we selected



Scheme 1 His and Arg based catalophore in MBH reaction.

cyclohexenone **2** and *p*-bromobenzaldehyde. All the versatile catalysts from a repository of MBH reaction[24] (DBU, DABCO, DBN, Et_3N, PPh_3, DMAP, His) resulted in 0–5% conversion to the MBH adduct **3l**. The product **3l** was formed in isolable amounts (22% conversion) only with imidazole. On the other hand, both lysozyme C **5c** and myoglobin **5e** resulted in 73% and 79% conversions (**3l**, Scheme 1).

The culmination of new enzymatic activity for non-natural transformations has relied on directed evolution by protein engineering. The focus has been to reconstruct the pre-existing protein domains. The key to success of the present report is hinged on the choice of catalophore and order of protein assembly. The latter is likely to provide an organic environment that can enhance the local concentration of substrates or render catalophore at the interface of two proteins or both. We have demonstrated that suitably spaced His and Arg residues can offer to serve as the catalytic site for efficient nucleophilic catalysis in MBH reaction. The results emphasize on controlling the oligomeric state of the protein involved in catalysis. This report will draw attention towards harvesting synergy of catalysis and protein self-assembly that can capably offer adaptive systems and expand the spectrum of biocatalysis.

Acknowledgements

P. N. J., L. P. and N. K. D. are recipient of research fellowship from CSIR and UGC, India. V. R. is a Ramanujan Fellow (SERB, India). The authors are thankful to SERB, DAE and DBT, India for research grants.

Notes and references

1 P. Grunwald, *Industrial Biocatalysis*, CRC Press, 2014.
2 (*a*) G. Kiss, N. Çelebi-Ölçüm, R. Moretti, D. Baker and K. N. Houk, *Angew. Chem., Int. Ed.*, 2013, **52**, 5700–5725; (*b*) X. Garrabou, T. Beck and D. Hilvert, *Angew. Chem., Int. Ed.*, 2015, **54**, 5609–5612.
3 W. P. Jencks, *Catalysis in Chemistry and Enzymology*, McGraw-Hill, New York, 1987.
4 A. Tramontano, K. D. Janda and R. A. Lerner, *Science*, 1986, **234**, 1566–1570.
5 H. Nam, N. E. Lewis, J. A. Lerman, D.-H. Lee, R. L. Chang, D. Kim and B. O. Palsson, *Science*, 2012, **337**, 1101–1104.
6 M. T. Reetz, *J. Am. Chem. Soc.*, 2013, **135**, 12480–12496.
7 (*a*) K. Fesko and M. Gruber-Khadjawi, *ChemCatChem*, 2013, **5**, 1248–1272; (*b*) E. Busto, V. Gotor-Fernandez and V. Gotor, *Chem. Soc. Rev.*, 2010, **39**, 4504–4523.
8 S. Bjelic, L. G. Nivon, N. Celebi-Olcum, G. Kiss, C. F. Rosewall, H. M. Lovick, E. L. Ingalls, J. L. Gallaher, J. Seetharaman, S. Lew, G. T. Montelione, J. F. Hunt, F. E. Michael, K. N. Houk and D. Baker, *ACS Chem. Biol.*, 2013, **8**, 749–757.
9 M. T. Reetz, R. Mondiere and J. D. Carballeira, *Tetrahedron Lett.*, 2007, **48**, 1679–1681.
10 M. Svedendahl, K. Hult and P. Berglund, *J. Am. Chem. Soc.*, 2005, **127**, 17988–17989.
11 (*a*) R. Robiette, V. K. Aggarwal and J. N. Harvey, *J. Am. Chem. Soc.*, 2007, **129**, 15513–15525; (*b*) V. K. Aggarwal, S. Y. Fulford and G. C. Lloyd-Jones, *Angew. Chem., Int. Ed.*, 2005, **44**, 1706–1708; (*c*) D. Roy and R. B. Sunoj, *Chem. - Eur. J.*, 2008, **14**, 10530–10534.
12 T.-W. Jeng and H. Fraenkel-Conrat, *FEBS Lett.*, 1978, **87**, 291–296.
13 If protein is replaced by a catalyst (imidazole) that would not form assemblies, it results in a linear increase in conversion with respect to the concentration.
14 H. Kim, R. C. Deonier and J. W. Williams, *Chem. Rev.*, 1977, **77**, 659–690.
15 L. J. Williams, L. Adcock-Downey and M. L. Pusey, *Biophys. J.*, 1996, **71**, 2123–2129.
16 Y.-C. Kim and A. S. Myerson, *J. Cryst. Growth*, 1994, **143**, 79–85.
17 A. Shukla, E. Mylonas, E. Di Cola, S. Finet, P. Timmins, T. Narayanan and D. I. Svergun, *Proc. Natl. Acad. Sci.*, 2008, **105**, 5075–5080.
18 H. H. Raeymaekers, H. Eisendrath, A. Verbeken, Y. van Haverbeke and R. N. Muller, *J. Magn. Reson.*, 1989, **85**, 421–425.
19 W. S. Price, F. Tsuchiya and Y. Arata, *J. Am. Chem. Soc.*, 1999, **121**, 11503–11512.
20 (*a*) M. Biancalana, K. Makabe, A. Koide and S. Koide, *J. Mol. Biol.*, 2009, **385**, 1052–1063; (*b*) M. Bhattacharya, N. Jain, P. Dogra, S. Samai and S. Mukhopadhyay, *J. Phys. Chem. Lett.*, 2013, **4**, 480–485; (*c*) A. Bhattacharya, R. Prajapati, S. Chatterjee and T. K. Mukherjee, *Langmuir*, 2014, **30**, 14894–14904; (*d*) U. Anand and M. Mukherjee, *Langmuir*, 2013, **29**, 2713–2721.
21 S. Chattoraj, A. K. Mandal and K. Bhattacharyya, *J. Chem. Phys.*, 2014, **140**, 115105.
22 U. Anand, C. Jash, R. K. Boddapalli, A. Shrivastava and S. Mukherjee, *J. Phys. Chem. B*, 2011, **115**, 6312–6320.
23 R. Singudas, S. R. Adusumalli, P. N. Joshi and V. Rai, *Chem. Commun.*, 2015, **51**, 473–476.
24 (*a*) D. Basavaiah, B. S. Reddy and S. S. Badsara, *Chem. Rev.*, 2010, **110**, 5447–5674; (*b*) K. Kaur and I. N. N. Namboothiri, *Chimia*, 2012, **66**, 913–920.