Supplementary information

An excess of catalytically required motions inhibits the scavenger decapping enzyme

Ancilla Neu1, Ursula Neu2,3, Anna-Lisa Fuchs1, Benjamin Schlager1,4, Remco Sprangers1

1 Max Planck Institute for Developmental Biology, Spemannstrasse 35, 72076 Tübingen, Germany
2 Interfaculty Institute of Biochemistry, University of Tübingen, Hoppe-Seyler-Strasse 4, 72076, Tübingen, Germany
3 Current address: The Francis Crick Institute, Mill Hill Laboratory, The Ridgeway, Mill Hill, London NW7 1AA, United Kingdom
4 Current address: FluidSolids AG, Zürich, Switzerland

*To whom correspondence should be addressed:
remco.sprangers@tuebingen.mpg.de
Supplementary Results

Supplementary Figure 1 | Structures of the human DcpS and spectra of the isolated yeast enzyme.

(a) Chemical structure of an mRNA with a 5’ protecting cap. The nomenclature of the sugars, base and phosphate backbone are indicated in red, blue and green respectively. The methyl group in the cap structure is indicated with a red sphere. The scavenger decapping enzyme hydrolyses the phosphate 5’ 5’ linkage at the position indicated with a red waved line for mRNA species with n < 10. For the m⁷GpppG substrate used in this study n=0 and the first base (base 1) is a guanine. The products of the hydrolysis of m⁷GpppG (m⁷GMP and GDP) are indicated.
(b) The human DcpS enzyme contains two protein chains that form a dimeric N-terminal lid domain and a dimeric C-terminal domain. The ligand free protein is symmetric (left)$^1$. In the substrate bound protein the lid domain undergoes a large conformational change (right)$^2$. In the structure of the human protein two substrates are bound to one dimeric enzyme (one in the open binding site and one in the closed binding site).

(c) Superposition of the structures of the human (1ST0)$^2$ and yeast (this study) scavenger decapping enzymes in complex with ligand. For clarity the ligands are not displayed in the ribbon diagram.

(d) Methyl TROSY spectra of the yeast full length enzyme (left panel) and the isolated N- and C- terminal domains (middle panels). A superposition of the spectra in the first three panels is shown on the right. The spectra of the isolated domains overlay very well with the spectrum of the full-length protein indicating that the two domains behave as independent units in the full-length protein.
Supplementary Figure 2 | Assignment of methyl groups in the yeast scavenger decapping enzyme.

(a) Example of the mutational approach for the methyl group assignments. Single Ile or Met residues were mutated into closely related amino acids. Spectra of the WT and mutant protein where then compared. An overlay of the spectra of the WT (black) and I311 (red) Dcs1p allows for the assignment of I311 in the symmetric (left) and asymmetric (right) conformation of the enzyme.

(b) Structure of the yeast Dcs1p enzyme (see also Figure 1), where all Isoleucine-δ1 and methionine methyl groups are shown as spheres. The assigned methyl groups are coloured orange.
Supplementary Figure 3a | Determination of the affinities of the first and second binding event for the interaction between WT (H268N) DcpS and m7GpppG substrate.

Top: Selected regions of NMR spectra during the NMR titration (regions are indicated with a box in Figure 2b and 2c). The m7GpppG:DcpS (dimer) ratios are indicated in the spectra. The spectra on the left report on the first binding event (that is in slow exchange on the NMR chemical shift timescale), the spectra on the right report on the second binding event (that is fast on the NMR chemical shift timescale).

Middle: Identical to Figure 2 of the main paper, shown here for completeness. Plot of the changes in the peak intensities for residues I12 and I36 (open and closed resonances) and the changes in the peak position for residues M153 (CSP in...
carbon) and I42 (CSP in protons) for 15 titration points (see a-c). All data points of the titration were simultaneously fitted to one sequential binding event (see below). The extracted kD values indicate that the first binding event takes place in the high nano-molar range and that the second binding event has a dissociation constant of 105 (±20) μM. The large uncertainty in the kD value for the first binding event results from the limited number of titration points that define the event. For clarity the scale of the x-axis is shown different below and above the 1:1 m7GpppG:DcpS ratio, as is indicated by the waveform.

Bottom: The ITC thermograph of the interaction between the yeast scavenger decapping enzyme (H268N) and m7GpppG substrate displays three events. One of these events could be a result of protein aggregation in the ITC cell (indicated with a red circle); note that no protein aggregation or precipitation has been observed in any of our NMR experiments. The red line in the right panel corresponds to the best fit of the data using a model of sequential binding. This third event prevents the extraction of accurate binding values for the first binding event (indicated with a red dotted box around the indicated extracted values). The kD for the second binding event extracted using NMR is in agreement with the one extracted from the ITC experiment. Note that n=1 corresponds to the stoichiometry of the first and of the second binding event, that both have an occupancy of 1.

In order to extract affinities from the ITC data, we performed baseline correction and peak integration using the NanoAnalyze program (TA instruments) and used in house written scripts to extract the associated binding affinities (the software provided by the manufactures did not include the sequential binding model).

In the case of sequential binding, a mixture of free protein, protein with one ligand bound and protein with two ligands bound can be formed. The associated fractions (F₀, F₁ or F₂) are given by:

\[
F_0 = \frac{1}{p} \quad F_1 = \frac{K_d^1 [L]}{p} \quad F_2 = \frac{K_d^1 K_d^2 [L]^2}{p}
\]
in which $[L]$ is the free ligand concentration and

$$P = 1 + K^1_a[L] + K^2_a K^3_a [L]^2$$

where $K^1_a$ and $K^2_a$ refer to the binding constants associated with the first and second binding event, respectively, that are defined according to:

$$K^1_a = \frac{[PL]}{[P][L]} \quad K^2_a = \frac{[PL_2]}{[P][L]}$$

$[P]$ is the concentration of free protein, $[PL]$ is the concentration of singly occupied protein and $[PL_2]$ is the concentration of double occupied protein. The concentration of free ligand $[L]$ is determined numerically and iteratively based on the binding constants and total ligand concentration $L_t$ at a given point during the titration, where

$$L_t = [L] + P_t * (F_1 + 2F_2)$$

and $P_t$ is the total protein concentration.

In an ITC experiment the heat content at a certain time during the titration is given by:

$$Q = P_t V_0 \left[ F_1 \Delta H_1 + F_2 (\Delta H_1 + \Delta H_2) \right]$$

where $V_0$ is the volume of the ITC cell and $\Delta H_1$ and $\Delta H_2$ are molar heats of ligand binding for the first and second binding steps respectively. In the fitting procedure, the parameters $\Delta H_1$, $\Delta H_2$, $K^1_a$, $K^2_a$ and $n$ (the total number of binding events per step) are optimized to minimize the difference between the observed and calculated heat content at all steps of the titration.
In the NMR titration experiments, the first binding event is slow on the chemical shift timescale. Thus, the intensity of the original peaks directly reports on the fraction of free protein ($F_0$), whereas the newly appearing peaks report on the sum of the singly and double occupied protein ($F_1+F_2$) (in case the second binding event does not perturb the position of the newly appearing resonance). The second binding event is fast on the NMR timescale and therefore, the peak position reports on the fraction of doubly bound protein $F_2$. To extract the affinities from the NMR titration experiments the peak intensities that report on the first binding event and the peak positions that report on the second binding event were simultaneously fitted. To that end the parameters $K_a^1$, $K_a^2$, $n$, the resonance position of the fully bound second binding site and a scaling factor for each of the two slow exchanging resonances (to correct for potential differences in line-width of the methyl resonances in the closed and open sites) were optimized to minimize the difference between the observed and calculated spectral properties at all steps of the titration.

Errors in all fits are based on 100 MC (Monte Carlo) simulations where the data points used for the fitting were randomly varied based on the uncertainties in the experimental data.
Supplementary Figure 3b | Determination of the affinities of the first and second binding event for the interaction between K126A (H268N) DcpS and m\(^7\)GpppG substrate.

Top: Regions of the NMR spectra during the NMR titration experiment. See legend to figure S3a.

Middle: Fit of the peak intensities and positions result in a kD of the first binding event that is in the high nano-molar range and a kD for the second binding event that is 196 (±31) \(\mu\)M.

Bottom: ITC thermogram for the interaction of Dcs1p (H268N, K126A) with m\(^7\)GpppG that displays a sequential interaction mode. The red line corresponds to the fit of the data; the extracted binding constants are indicated next to the graph. Please note that the third event that we observed in the ITC experiments
with the WT protein is not observed in the ITC experiment using the K126A mutant. As a result, the kD for the first binding event can be extracted from the ITC data for the K126A mutant.

It is important to note that the affinities of m7GpppG for the WT (H268N) and the K126A (H268N) enzymes are very similar. In addition, the NMR spectra during the titration of the substrate to both forms of the enzyme show exactly the same behaviour (compare Figures S3a and S3b). This establishes that both proteins use the same sequential binding mechanism and the K126A mutation does not change the mechanism of the enzyme.
Supplementary Figure 3c | Analysis of the interaction between the C-terminal domain of Dcs1p and m7GpppG.

Top: ITC thermogram for the interaction of the C-terminal domain of Dcs1p with m7GpppG substrate. The affinity of the substrate to the C-terminal domain is comparable to the affinity of the substrate to the open (second) binding site in the full-length enzyme. This confirms that the interaction of the substrate in the open binding site is mainly mediated through the C-terminal domain.

Bottom left: Methyl TROSY NMR spectrum of the C-terminal domain of Dcs1p in the absence (red) and presence (blue) of a 5-fold excess of m7GpppG. The same residues experience chemical shift changes in the isolated domain and in the full-length protein (Figure 2c), indicating that the m7GpppG substrate interacts in the same manner with the isolated C-terminal domain and the full length protein.

Bottom right: To judge how the substrate structurally interacts with the open binding pocket of the yeast protein, we modeled the m7GpppG substrate into the empty open binding pocket of our yeast structure. The model is based the human Dcs1p enzyme (1ST0) that contains substrate in the open and closed binding sites. The N-terminal domain is colored in blue, the hinge region in red and the C-
terminal domain is shown in green, the ligand m7GpppG in yellow. In our model, the distance between the Cδ1 methyl group of I42 (pink sphere) and the substrate (dashed line) is between 12 and 17 Å (depending on how the substrate is exactly modeled). The shift of I42 in our NMR titration experiments can thus not be due to direct contacts between the substrate and the methyl group in the N-terminal domain. Residue K126 that alters the dynamics of the protein is located in the hinge region, in-between the bound second substrate and I42 in the N-terminal domain.
Supplementary Figure 4 | Asymmetry of the enzyme is required for activity.

(a) The Y94A Dcs1p enzyme interacts with m^7GpppG (left spectrum). The WT Dcs1p enzyme interacts with (non-methylated) GpppG (two right spectra: Ile, respectively Met region). In both cases, the enzyme fails to adopt an asymmetric conformation. This can be clearly seen from e.g. residues I12 and I36, that don’t experience peak splitting as we observed for the WT protein in the presence of m^7GpppG (Figure 2B). It should be noted that the affinity of the protein for the ligand is unaffected by the mutations as the extent of the chemical shift changes observed here is comparable to the chemical shift changes that are observed for the open binding site in WT protein with m^7GpppG (compare e.g. M153 in Figure S3a).

(b) ^31P spectra of the m^7GpppG or GpppG in the presence of the decapping enzyme. Upon hydrolysis, the phosphate spectrum of the ligand changes. m^7GpppG is not hydrolysed by the Y94A enzyme (left) and GpppG is not hydrolysed by WT Dcs1p (middle). As a reference the spectra of m^7GpppG in the presence of WT enzyme are shown on the right, where the substrate is hydrolysed completely within 14 hours.
Supplementary Figure 5 | Substrate competes the product out of the enzyme.

(a) DcpS without and with the product m\textsuperscript{7}GMP. In the presence of m\textsuperscript{7}GMP the enzyme adopt an assymmetric conformation as can be judge from the splitting of the resonances of I12 and I36. Note that I12 appears at lower carbon chemical shifts compared the other spectra shown in this paper due to a difference in the aliasing of the spectrum.

(b) DcpS binds the m\textsuperscript{7}GMP product with an affinity of 1.05 ± 0.64 \(\mu\)M in the first binding event. A second binding event is not observed, indicating that the product has no affinity for the open binding site. Using NMR spectroscopy and ITC we have not been able to detect any interaction of the GDP product with either the first or the second binding site (data not shown), underscoring the importance that the 7-methyl group plays in the recognition process between the nucleotides and the enzyme (see Figure S4).

(c) The substrate is able to compete the product out of the closed binding site. The NMR spectra of the WT protein bound to m\textsuperscript{7}GMP and bound to m\textsuperscript{7}GpppG are very similar and in order to be able to clearly distinguish if substrate or product is bound in the closed binding pocket, we introduced a reporter isoleucine into
the enzyme (Leucine 279 to Isoleucine). The enzyme containing this reporter isoleucine was converted into an asymmetric conformation by the addition of the product \( m^7 \text{GMP} \). Subsequently, we added \( m^7 \text{GpppG} \) substrate in a stepwise manner. During the addition of the substrate the reporter isoleucine shifts from the \( m^7 \text{GMP} \) bound position to the \( m^7 \text{GpppG} \) bound position. This indicates that the asymmetric product bound form of the enzyme is transferred into an asymmetric substrate bound form. In other words, interaction of substrate in the second binding site results in an opening of the first, product filled binding site. As a result the product can rapidly leave the enzyme, as this exhibits no detectable interaction with the open binding pocket.
Supplementary Figure 6 | Cartoon representation of the K126A mutant enzyme.

The K126A mutation decouples the occupation of the second binding site from domain flipping motions. Unproductive motions are reduced or absent (horizontal dashed line) and catalytic turnover increases (curved dashed arrows). The K126A mutation is indicated with a cross.
**Supplementary Figure 7 | Activity assays under reducing and oxidizing condition.**

The activity of Dcs1p is not influenced by the formation of a disulphate bond in the N-terminal domain (that we observe in the crystal structure). Shown are three different time points from a degradation series in the absence (top) or presence (bottom) of 1 mM DTT (in the bottom spectra 1 mM DTT has been present during the complete purification). The spectra are identical within the noise, proving that the activity is independent of the formed disulphate bond. In both spectra 200 nM WT Dcs1p was mixed with 0.5 mM m^7GpppG substrate at 20°C.
## Supplementary Table 1 | Data collection and refinement statistics

### Data collection

| Parameter                  | Value                  |
|----------------------------|------------------------|
| Space group                | 19                     |
| Cell dimensions            |                        |
| \(a, b, c\) (Å)            | 87.99, 104.52, 189.96  |
| \(\alpha, \beta, \gamma\) (°) | 90.0, 90.0, 90.0       |
| Resolution (Å)             | 20(2.25)               |
| \(R_{\text{merge}}\)      | 6.7 (71.6)             |
| \(I / sI\)                | 12.04(2.35)            |
| Completeness (%)           | 99.5 (98.9)            |
| Redundancy                 | 3.3(3.3)               |

### Refinement

| Parameter                  | Value                  |
|----------------------------|------------------------|
| Resolution (Å)             | 2.25                   |
| No. reflections            | 83812                  |
| \(R_{\text{work}} / R_{\text{free}}\) | 0.235/0.262           |
| No. atoms                  |                        |
| Protein                    | 10323                  |
| Ligand/ion                 | 93                     |
| Water                      |                        |
| \(B\)-factors              |                        |
| Protein                    | 51.16                  |
| Ligand/ion                 | 48.64                  |
| Water                      |                        |
| R.m.s. deviations          |                        |
| Bond lengths (Å)           | 0.01                   |
| Bond angles (°)            | 1.35                   |

Values in parentheses are for highest-resolution shell.
Supplementary References

1. Chen, N., Walsh, M.A., Liu, Y., Parker, R. & Song, H. Crystal structures of human DcpS in ligand-free and m7GDP-bound forms suggest a dynamic mechanism for scavenger mRNA decapping. *J Mol Biol* **347**, 707-18 (2005).

2. Gu, M. et al. Insights into the structure, mechanism, and regulation of scavenger mRNA decapping activity. *Mol Cell* **14**, 67-80 (2004).