JMB

ELSEVIER

Enhanced Catalytic Efficiency of Aminoglycoside Phosphotransferase (3′)-IIa Achieved Through Protein Fragmentation and Reassembly

David E. Paschon, Zarana S. Patel and Marc Ostermeier*

Department of Chemical and Biomolecular Engineering Johns Hopkins University 3400 N. Charles St., Baltimore MD 21218, USA

Many monomeric proteins can be split into two fragments, yet the two fragments can associate to make an active heterodimer. However, for most locations in a protein such a conversion is not feasible, presumably due to inefficient assembly or improper folding of the fragments. For some locations, this can be overcome by fusion of the fragments to dimerization domains that facilitate correct assembly. A variety of heterodimers of aminoglycoside phosphotransferase (3′)-IIa (Neo) were created in which the Neo fragments required fusion to a pair of leucine zippers for activity *in vivo*. However, the ability of these heterodimers to confer kanamycin resistance to *Escherichia coli* cells was impaired compared to wild-type Neo, primarily due to poor production of soluble protein. The mutations R177S and V198E restored the kanamycin resistance to wild-type levels while maintaining the dependence on leucine zippers for activity. These mutations restored high levels of kanamycin resistance not through an improvement in the production of soluble protein but rather by conferring a large improvement in k_{cat}/K_m, surpassing that of Neo. Furthermore, whereas R177S and V198E served to improve k_{cat}/K_m 60-fold in the context of the heterodimer, the same mutations in the context of wild-type Neo had a ninefold negative effect on k_{cat}/K_m. This demonstrates the possibility that enzymes with improved catalytic properties can be created through a process involving fragmentation and fusion to domains that facilitate assembly of the fragments.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: protein fragment complementation; aminoglycoside phosphotransferase; protein engineering; protein complementation assay

*Corresponding author

Introduction

Protein fragment complementation occurs when a protein consisting of a single polypeptide chain is fragmented into two polypeptide chains, either by proteolytic or genetic methods, and the two fragments reassemble into a functional protein.[1,2] Although a variety of proteins have been shown to tolerate fragmentation at select sites, functional fragmentation sites are relatively rare.[3] However, if two fragments that do not appreciably assemble into an active protein are each fused to domains that

dimerize, the dimerization domains can assist the functional reassembly of the two fragments. Such a "protein-assisted reassembly" strategy has been used in the creation of protein-fragment complementation assays (PCAs), two-hybrid-like systems for identifying and interrogating protein–protein interactions.[4–11] For example, the enzyme β-lactamase has been bisected into two fragments that do not appreciably assemble into an active enzyme unless they are each fused to proteins that associate.[12,13]

An important goal of synthetic biology is the ability to modulate the function of biological molecules and pathways. Several PCAs have been demonstrated in which the dimerization domains depend on a small molecule for assembly, usually rapamycin-induced dimerization of FRB and FKBP.[6,7,13–17] Such small molecule-mediated PCAs are analogous to a strategy for the control of

protein function using chemical inducers of dimerization (CIDs). This strategy utilizes a CID to induce proximity of two proteins (A and B) through facilitating the dimerization of domains attached to proteins A and B.[18] If proteins A and B require proximity to perform a function, then the CID can be used to control that function in a dose-dependent manner. Control of a monomeric protein might be achieved by an analogous method if the protein is first fragmented into two fragments that require fusion to dimerization domains for assembly and these domains require a small molecule (e.g. CID) for dimerization.

We were interested in a number of issues related to the design and applicability of such a strategy for control of protein function. These issues included: (1) whether a systematic method could be used to identify fragments of a protein that would require fusion to dimerization domains for activity; (2) the utility of the linker domain between the fragmented protein and the dimerization domain; and (3) whether fragmented and reassembled proteins with activities equivalent to the unfragmented parental protein could be created. The importance of these issues stems, in part, from differences in design considerations for control of protein function and PCAs. The purpose of PCAs is not to control protein function but rather to use the function of the bisected protein as a readout for protein–protein interactions. Since PCAs must accommodate a variety of association geometries, the proteins to be tested for association and the fragments of the split protein are connected through long peptide linkers. In addition, the specific activity of the assembled protein is not of paramount concern for PCAs, since overexpression can compensate for poor specific activity and all that is required is a selectable or screenable phenotype *in vivo* when protein interactions occur.

Tafelmeyer *et al.* recently demonstrated a combinatorial method for generating libraries of split proteins that involved the circular permutation of DNA and homologous recombination in yeast.[10] They applied this strategy to identify heterodimers of the Trp1p protein for use in PCAs. Here, we demonstrate that highly active heterodimers of aminoglycoside phosphotransferase (3′)-IIa that require fusion to antiparallel leucine zippers for functional reassembly can be identified systematically using a combination of incremental truncation[3] and random mutagenesis. One of these heterodimer's catalytic activity exceeded that of the parental enzyme by 18-fold. Characterization of this heterodimer demonstrated that the effects of fragmentation and point mutagenesis can be highly non-additive and, quite surprisingly, that fragmentation and reassembly alone can result in significant (170-fold) improvements in enzymatic activity.

Results

Aminoglycoside phosphotransferase (3′)-IIa

(commonly known as the neomycin-resistance protein or Neo) was chosen as the model protein for our studies. Neo inactivates aminoglycoside antibiotics such as kanamycin by phosphoryl transfer from ATP to the 3′ hydroxyl group. Neo offered an easy selection for active heterodimers (plating on medium containing kanamycin) and one that is readily amenable to increased selection stringency by increasing the concentration of kanamycin. In addition, any heterodimers of Neo that required protein domains to assist their assembly would have potential for use in PCAs in both bacteria and eukaryotic cells.[4]

We utilized a combinatorial approach for identifying fragment pairs of Neo that could assemble into an active enzyme.[3,19] The first step in this process is to select large overlapping fragments of Neo that coded for inactive fragments (Figure 1). Based on the crystal structure of a closely related homolog,[20] and proposed conserved active site
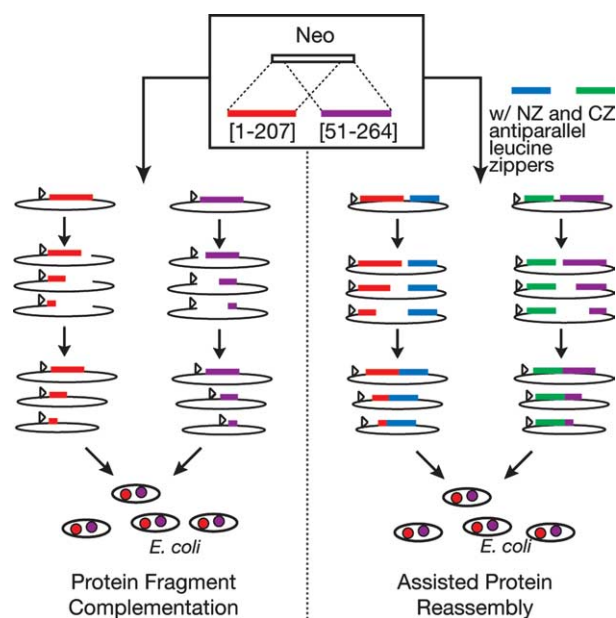

Figure 1. A depiction of the creation of incremental truncation libraries of Neo. Large, overlapping gene fragments coding for inactive N-terminal and C-terminal fragments of Neo were cloned into vectors designed for performing incremental truncation.[19] The gene fragment for the N-terminal fragment was truncated in the 3′ to 5′ direction and the gene fragment coding for the C-terminal fragment was truncated in the 5′ to 3′ direction. For libraries constructed for identifying fragments that exhibit protein fragment complementation (left side), N-terminal gene fragments were fused to a series of stop codons in all three frames and C-terminal gene fragments were fused to an ATG start codon. Individual N-terminal and C-terminal truncation libraries were first constructed and then cotransformed into *E. coli* cells. For libraries constructed for identifying fragments that can assemble when fused to leucine zippers (right side), the same procedure was followed, except that the vectors contain genes coding for the NZ and CZ leucine zippers[5] (with or without segments coding for peptide linkers), which were fused to the gene fragments on the truncated end of the gene.

residues,[21] we conservatively chose as our largest N-terminal fragment the DNA coding for amino acid residues 1–207 (*neo[1-207]*) and as our largest C-terminal fragment the DNA coding for amino acid residues 51–264 (*neo[51-264]*). Thus, our libraries would be searching the overlapping region between these two fragments (51–207) for active fragment pairs. We confirmed that individually neither of the starting gene fragments was able to confer resistance to kanamycin above background levels.

Construction of incremental truncation libraries

A series of incremental truncation libraries (Figure 1) were constructed in order to systematically identify fragment pairs of Neo that could assemble into an active enzyme, either with or without leucine zippers NZ and CZ (Figure 2(a)). NZ and CZ dimerize in an antiparallel orientation and have been shown to assist the reassembly of fragments of GFP.[5] The purpose of one library pair was to search for fragment pairs of Neo that would assemble in an unassisted fashion without fusion to NZ and CZ (left side of Figure 1). In these libraries, 3′ truncations of *neo[1-207]* were fused to a cassette of stop codons in all three reading frames and 5′-truncations of *neo[51-264]* were fused to an ATG start codon. Libraries with antiparallel leucine zippers NZ and CZ (right side of Figure 1) were constructed such that 3′ truncations of *neo[1-207]* were fused to DNA coding for NZ and 5′-truncations of *neo[51-264]* were fused to DNA coding for CZ. Fusions between leucine zippers and truncations of *neo* were either (i) without a region coding for a linker, (ii) with a region coding for a GSGG linker (leucine zippers NZ1 and CZ1), or (iii) with a region coding for a (GSGG)$_2$ linker at the fusion site (leucine zippers NZ2 and CZ2). In addition, analogous libraries in which fragments of *neo* were fused to sequences coding for parallel zippers P-NZ and P-CZ *via* a (GSGG)$_3$ were constructed. The sequence of these parallel zippers was based on previously designed parallel zippers.[22] Individual truncation libraries of $>10^6$ transformants each were created. For each library, a suitable distribution of truncations was verified by PCR analysis of randomly selected library members. Plasmid DNA from appropriate N and C-terminal library pairs were co-transformed into *Escherichia coli* DH5α-E cells and stored in aliquots at −80 °C.

Selection of active heterodimers

Functional heterodimers were selected on LB plates containing kanamycin. DH5α-E cells without a plasmid, with pDIM-N2-Neo[1-207], or with pDIM-C8-Neo[51-264] could not grow above 2.5 μg/ml of kanamycin. DH5α-E cells harboring pDIM-N2-Neo (containing the full-length *neo* gene) could grow at 500 μg/ml of kanamycin. Aliquots of the frozen, co-transformed libraries were thawed, diluted and plated at 37 °C on increasing concentrations of kanamycin. It was determined (by DNA sequencing) that some of the colonies that grew were the result of recombination between the two plasmids to restore a full-length *neo* gene in a recombined plasmid that was larger than either of the original two starting plasmids. Thus, any library member that grew on ≥ 5 μg/ml of kanamycin was further screened for the presence of a large plasmid.

(a)

```
NZ2  n-SGSGGGSGGAQLKKELQANKKELAQLKWELQALKKELAQ-c
 NZ1  n-SGSGGAQLKKELQANKKELAQLKWELQALKKELAQ-c
  NZ  n-SAQLKKELQANKKELAQLKWELQALKKELAQ-c
              |X|  |X|  |X|  |X|
  CZ  c-QALKKELAQNKWELQALKKELAQLKKELQASAM-n
 CZ1  c-GGSGQALKKELAQNKWELQALKKELAQLKKELQASAM-n
CZ2  c-GGSGGGSGQALKKELAQNKWELQALKKELAQLKKELQASAM-n


P-NZ  n-SGSGGGSGGGSGGAQLEKELQALEKENAQLEWELQALEKELAQ-c
            |X|  |X|  |X|  |X|
P-CZ      n-MASAQLKKKLQALKKKNAQLKWKLQALKKKLAQGSGGGSGGGSGG-c
```

(b)

```
  CZ1  c-GGSGQALKKELAQNKWELQALKKELAQLKKELQASAM-n
            |X|  |X|  |X|  |X|
NZ1mut  n-SGSGGAQLKKELQANEKELAQLKWKLQALKKKLAQ-c
            |X|  |X|  |X|  |X|
CZ1mut  c-GGSGQALKKELAQNKWKLQALEKELAQLEKELQASAM-n
            |X|  |X|  |X|  |X|
 NZ1  n-SGSGGAQLKKELQANKKELAQLKWELQALKKELAQ-c
```
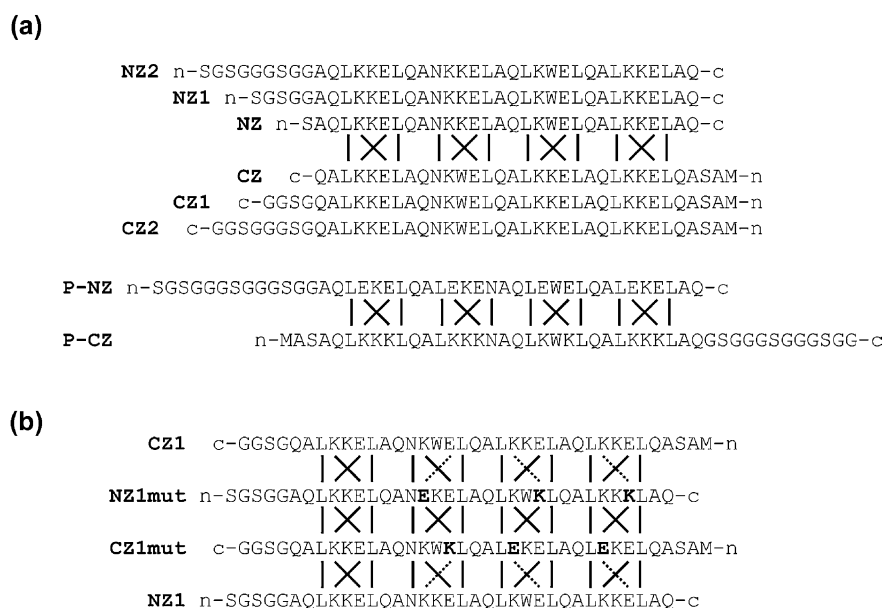
Figure 2. Sequences of leucine zippers NZ and CZ[5] and variants thereof. Key positive hydrophobic and charge–charge interactions are shown (continuous lines). (a) Leucine zippers used in the construction of incremental truncation libraries. (b) Sequences of leucine zippers NZ1mut and CZ1mut. Mutations are shown in bold and negative charge–charge interactions are shown by the broken lines.

Any library member containing an unexpectedly large plasmid was designated as a recombinant and discarded. For the library members that grew on ≥5 μg/ml of kanamycin and bore two plasmids in the expected size range, the individual plasmids were isolated and (i) transformed separately into DH5α-E to confirm that neither plasmid alone conferred kanamycin resistance and (ii) cotransformed into DH5α-E to confirm that only the plasmid pair conferred kanamycin resistance.

The frequencies of active heterodimers in the libraries are shown in Figure 3(a). No fragment pairs from libraries of Neo fragments without leucine zippers were found that conferred kanamycin resistance. This indicates that Neo is not amenable to protein fragment complementation in the region between residues 51 and 207. However, libraries in which the fragments of Neo were fused

Figure 3. Active heterodimers of Neo. (a) Frequencies in incremental truncation libraries of heterodimers capable of conferring growth at the indicated concentrations of kanamycin at 37 °C. Libraries contained fragments of Neo fused to antiparallel leucine zippers or parallel leucine zippers, with or without a linker sequence. Black bars, antiparallel zippers and no linker; dark grey bars, antiparallel zippers and GSGG linker; light grey bars, antiparallel zippers and $(GSGG)_2$ linker; white bars, parallel zippers and $(GSGG)_3$ linker. (b) Sequences of randomly selected heterodimers. The diagonal broken line represents fragment pairs in which there is neither fragment overlap nor deleted residues. Fragment pairs above the diagonal line contain sequences common to both fragments (overlapping sequences). Fragment pairs within or near quadrant A (bound by two broken lines) could be considered to be at the domain boundary between the two structural domains of Neo.[23]

to antiparallel leucine zippers contained heterodimers capable of conferring growth on plates containing 5 μg/ml and 15 μg/ml, but not 50 μg/ml of kanamycin. The frequency of active heterodimers was considerably higher for libraries without a linker or with the GSGG linker than with the $(GSGG)_2$ linker. The frequency of active heterodimers with parallel leucine zippers and $(GSGG)_3$ linkers also was low.

Randomly selected active heterodimers lacking linkers were sequenced. The sequences of complementing fragments without linkers roughly clustered in different regions of the space searched (Figure 3(b)) with a bisection point near residues 50–60 predominating. Neo is folded into two continuous structural domains, with the domain boundary at residue 96.[23] Only a few fragment pairs (heterodimers 5ZU37-33 and 5ZI37-2 in Table 1) were found that could be considered to be near this boundary (quadrant A of Figure 3(b)). A few selected library members of the other libraries (i.e. those with linker-containing constructs with antiparallel zippers and those with parallel leucine zippers) were also sequenced, and these sequences were similar to those found without linkers. PCR analysis of additional linker-containing active heterodimers indicated that inclusion of linkers did not produce complementing fragments significantly different from those found without the use of a linker. Table 1 summarizes all the sequenced complementing fragments.
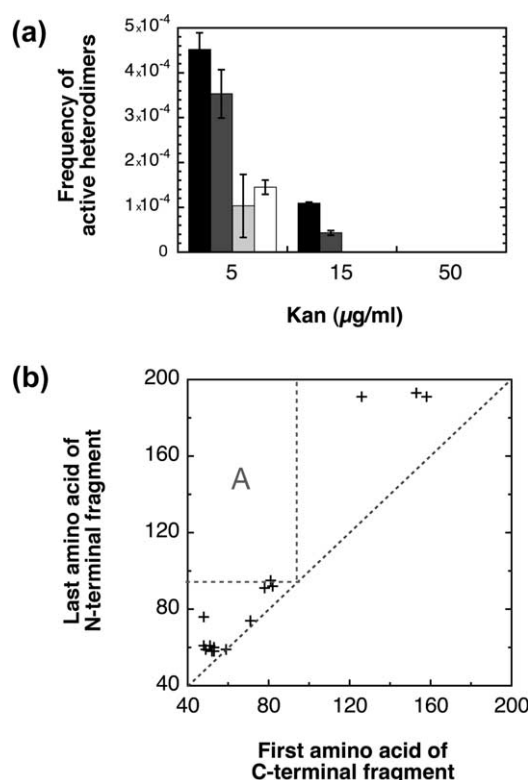
Characterization of heterodimer H1

Although active heterodimers were identified, they conferred a level of kanamycin resistance that

Table 1. Amino acid sequences of Neo heterodimers

Zippers	Heterodimer	N-terminal sequence	C-terminal sequence
NZ/CZ	5ZU37-45	1-58	53-264
	15ZI37-17	1-59	RS[51-264]
	15ZI37-19	1-60	53-264
	5ZU37-6	1-61	LRS[51-264]
	5ZU37-9	1-74	71-264
	5ZI37-19	1-76	LRS[51-264]
	5ZU37-33	1-91	78-264
	5ZI37-2	1-95	81-264
	5ZU37-38	1-191	158-264
	5ZU37-15	1-193	153-264
NZ1/CZ1	5ZLU37-3	1-58	52-264
	H1	1-59	59-264
	5ZLU37-5	1-92	82-264
	5ZLU37-10	1-191	126-264
NZ2/CZ2	5Z2LU37-3	1-68	LRS[51-264]
P-NZ/C-NZ	5PZU73-8	1-185	166-264
	5PZU73-1	1-57	59-264

Amino acid sequences are based on DNA sequencing. Amino acid numbers in Neo are indicated. For some C-terminal fragments, extra amino acid residues are present at the interface between the zipper/linker and the Neo fragment. These sequences arise from DNA in which truncation did not progress into the gene fragment but stopped in a non-coding region between the starting point for truncation in the plasmid and the gene fragment.

Table 2. Summary of kinetic constants for purified enzymes

Enzyme	k_{cat} (s^{-1})	$K_{m,Kan}$ (µM)	$K_{m,ATP}$ (µM)	$k_{cat}/K_{m,Kan}$ (s^{-1} µM^{-1})	$K_{i,Tobramycin}$ (µM)
Neo[a]	0.63	5.1±0.4	48±7	0.12	–
Neo[b]	4±0.8	3±0.6	45±9	1.3±0.4	–
Neo[c]	1.1±0.2	4.5±0.5	27±4	0.24±0.05	1.0±0.1
H1[d]	0.23±0.05	3.2±0.4	55±5	0.072±0.018	0.98±0.14
H1[e]	0.35±0.07	3.2±0.4	17±1	0.11±0.03	–
Ev1a[d]	0.12±0.02	0.028±0.005	27±1	4.3±1.0	0.047±0.001
Ev1a[e]	0.14±0.03	0.05±0.01	21±1	2.8±0.8	0.055±0.002
Neo(Ev1a)[c]	0.253±0.05	9.4±0.9	56±2	0.026±0.005	3.0±0.5

[a] From Perlin *et al.*[25]
[b] From Siregar *et al.*[24]
[c] This study. A His$_6$ tag is attached to the N terminus.
[d] This study. A His$_6$ tag is attached to the N terminus of the N-terminal fragment.
[e] This study. A His$_6$ tag is attached to the N terminus of the C-terminal fragment.

was over tenfold lower than that conferred by full-length Neo. Plausible reasons for the decreased kanamycin resistance conferred by the hetero-dimers included decreased specific activity and decreased production of active enzyme. To evaluate these hypotheses, studies were performed comparing the production and kinetic properties of wild-type Neo and heterodimer pair H1 (Neo[1-59]-GSGG-NZ and CZ-GSGG-Neo[59-264]). H1 was chosen as representative of the cluster of hetero-dimers split around residues 50–60.

Analysis by ELISA on fractionated cell lysates of cells grown under the conditions of selection (no induction and 37 °C) indicated that the amount of soluble H1 that is produced is 4.3(±0.4)% of the amount of soluble Neo that is produced. Western blot analysis showed that the ratio of soluble to insoluble Neo was approximately 1:1, whereas the ratio of soluble to insoluble H1 was approximately 1:20. Thus, one factor contributing to the decreased kanamycin resistance conferred by H1 is decreased production, presumably owing to increased aggregation. However, another possibility is a deficiency in transcription of the C-terminal fragment. Full-length Neo is produced from the same plasmid and *lac* promoter as the N-terminal fragment, whereas the C-terminal fragment is produced from a plasmid with a lower copy number and a *tac* promoter.

The kinetic constants of purified Neo and H1 (Table 2) were determined using a fluorescent assay that couples ADP generation and NADH oxidation.[24] In order to facilitate purification, a His$_6$ tag was added to the N terminus of wild-type Neo and the N-terminal fragment of H1. Separate constructs were created in which the His$_6$ tag was added to the N terminus of the C-terminal fragments of H1. The enzymes were purified by nickel affinity chromatography and determined to be ≥95% pure when visualized on SDS/polyacrylamide gels stained with Coomassie brilliant blue. It was confirmed that neither purified protein contained any activity that interfered with the assay (i.e. an enzyme capable of ATP hydrolysis or NADH oxidation) by performing the assay in the absence of kanamycin. The kinetic parameters obtained for wild-type Neo ($k_{cat}=1.1$ s^{-1}, $K_{m,kan}=4.5$ µM and $K_{m,ATP}=27$ µM) were within the range of previous studies on non-His-tagged Neo[24,25] (see Table 2), indicating that the N-terminal His-tag does not significantly affect catalytic activity. For the characterization of H1, the protein concentration used during the assay (60 nM) was well above the dissociation constant of the two fragments, since the heterodimers' specific activity did not decrease with protein concentration over the protein concentration range of 3–120 nM (Figure 4). Fragmentation of Neo and fusion to leucine zippers to create H1 resulted in a fairly active enzyme with only a fivefold decrease in k_{cat} and a threefold decrease in $k_{cat}/K_{m,kan}$. The kinetic parameters of H1 were essentially independent of whether the His tag was attached to the N-terminal fragment or to the C-terminal fragment. Thus, the major reason for decreased resistance to kanamycin in cells expressing H1 is the 23-fold decrease in production of soluble protein.
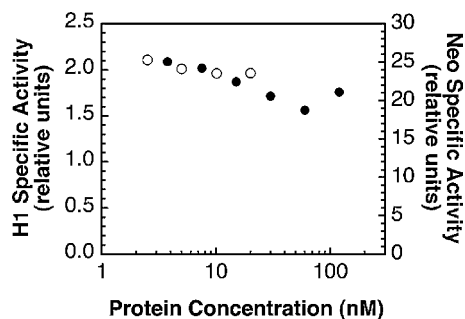

Figure 4. Specific activity of H1 does not decrease with protein concentration over the protein concentration range of 3–120 nM. The specific activities of H1 (filled circles) and Neo (open circles) at the indicated concentrations of enzyme were defined as the initial rate of aminoglycoside phosphorylation at 25 °C (as measured by the coupled spectrofluorometric assay using 300 µM kanamycin and 150 µM ATP) divided by the protein concentration.

Table 3. Effect of dimerization domain identity on kanamycin resistance

C-terminal Neo fragment	C-terminal dimerization domain	MIC kanamycin (µg/ml)[a] N-terminal dimerization domain[b]	
		NZ	NZmut
H1	CZ	100	100
	Czmut	25	100
Ev1a	CZ	800	400
	Czmut	200	400

[a] Minimum inhibitory concentration of kanamycin on LB plates at 37 °C; concentrations of kanamycin tested were 0, 2.5, 5, 10, 25, 50, 100, 200, 400, 800 and 1600 µg/ml.
[b] The N-terminal Neo fragments of H1 and Ev1a have identical sequences.

Role of leucine zipper dimerization in assembling the H1 fragments of Neo

Variants of H1 in which either one or both of the leucine zippers were removed failed to confer kanamycin resistance above background levels. However, this showed only that both leucine zippers are required to confer kanamycin resistance to *E. coli*. The function of the leucine zippers may be to improve stability or to make the fragments less sensitive to protease rather than to assist in the assembly of the fragments *via* dimerization. In order to test if NZ1-CZ1 dimerization was required, we constructed a variant of CZ1 (CZ1mut; Figure 2(b)) containing three mutations that lowers the affinity about 150-fold by the introduction of three charge–charge repulsions.[26] This lowered the MIC fourfold (Table 3). When compensating mutations designed to restore the correct charge–charge interactions were introduced into NZ1 (to make NZ1mut; Figure 2(b)), the MIC was restored to the level observed when using NZ1 and CZ1. This implicates leucine zipper dimerization as at least a partial requirement for H1 fragment assembly. Constructs containing NZ1mut/CZ1 (i.e. the opposite charge–charge repulsions of NZ1/CZ1mut) conferred no decrease in MIC

(Table 3). However, the effect of mutations in NZ on affinity for CZ is not known, and some triple mutations that introduced three charge–charge repulsions in CZ only had a twofold effect on affinity for NZ.[26] Thus, it is possible that the NZ1mut/CZ1 does not have the expected decrease in affinity. The complication in using NZ1mut/CZ1mut to address the role of NZ1/CZ1 dimerization in assembly arises from the fact that the pairs are not completely orthogonal.

Random mutagenesis and selection for improved heterodimers

Heterodimer pair H1 was selected to undergo random mutagenesis and selection for conferring improved kanamycin resistance. H1 conferred DH5α-E with the ability to grow on LB plates containing 50 µg/ml but not 100 µg/ml kanamycin when plated from stationary cells grown in liquid culture (a less stringent condition than plating from frozen stocks, as was done in the original selection when H1 was found to grow at 15 µg/ml of kanamycin). Error-prone PCR was performed as described,[27] except that the concentration of manganese chloride was adjusted such that ∼60% of the library should receive one or two non-synonymous mutations.[28] Individual libraries of $>1.1 \times 10^5$ each were created and subsequently co-transformed to create a library of 3.4×10^6 transformants. A total 1.1×10^6 co-transformants were plated onto plates containing 500 µg/ml of kanamycin. Of the 29 colonies that grew at 500 µg/ml of kanamycin, 22 were found to be false-positives bearing a full-length Neo gene. Of the seven remaining colonies, sequencing revealed four distinct fragment pairs (Table 4). No mutation was found in the region coding for the linker or the leucine zippers. For all four, the heterodimeric nature and ability to grow at 500 µg/ml of kanamycin was confirmed by retransformation as before. Interestingly, ELISA analysis indicated that the production level of soluble heterodimers of the four evolved heterodimers was not significantly different from H1. Compared to wild-type Neo production (set to a value of 100), the production of H1 (4.3±0.4), Ev1 (5.3±0.1), Ev2 (4.6±0.3), Ev3

Table 4. Kanamycin resistance conferred by heterodimers using NZ1 and CZ1 dimerization domains

Enzyme	N-terminal mutations	C-terminal mutations	MIC[a] kanamycin (µg/ml)
H1			100
Ev1	N58S	R177S, V198E	800
Ev2	C31R	M120L	800
Ev3	C31R	K175E, V198E	800
Ev4	C31R, D52Q	D118E, Q155L	800
H1(N58S)	N58S		100
H1(R177S)		R177S	200
H1(V198E)		V198E	200
Ev1a		R177S, V198E	800

[a] Minimum inhibitory concentration of kanamycin on LB plates at 37 °C; concentrations of kanamycin tested were 0, 2.5, 5, 10, 25, 50, 100, 200, 400, 800, and 1600 µg/ml.

(5.5 ± 0.2) and Ev4 (2.8 ± 0.2) were all poor (the error is the standard deviation from three independent cultures). This suggested strongly that the improved level of antibiotic resistance conferred to cells by the mutations of Ev1 resulted, to a large extent, from improved catalytic activity.

In vivo characterization of Ev1

Ev1 was selected for further characterization. First, it was confirmed that Ev1 still required fusion to leucine zippers for activity. Variants of Ev1 in which either one or both of the leucine zippers were removed failed to confer kanamycin resistance above background levels. In order to evaluate qualitatively the relative effect of the three mutations (N58S, R177S, and V198E), variants of H1 were constructed with only one of the mutations and cells producing the single mutants were plated on increasing amounts of kanamycin. R177S and V198E separately resulted in an improved MIC for kanamycin (compared to H1), but N58S did not (Table 4). The combination of R177S and V198E resulted in the same kanamycin resistance as Ev1, and the heterodimer bearing these two mutations was designated Ev1a. These mutations are neither at the interface of the two fragments nor proximal to the active site (Figure 5(a)).

Characterization of the evolved heterodimer Ev1a

Variants of Ev1a in which one or both of the leucine zippers were replaced with the mutated leucine zippers (i.e. CZmut and NZmut) conferred a pattern of antibiotic resistance that was similar to that conferred by the analogous variants of H1 (Table 3). Replacing CZ1 with CZ1mut lowered the MIC fourfold, an effect that could be overcome partially by compensating for the CZ1mut mutations by replacing NZ1 with NZ1mut.

Since mutations of Ev1a did not improve the production of soluble protein significantly, the Ev1a mutations must result in an enzyme with improved kinetic parameters. The R177S and V198E mutations resulted in a dramatic 160-fold decrease in $K_{m,kan}$, resulting in a 60-fold improvement in $k_{cat}/K_{m,kan}$ (Table 2). The catalytic efficiency of Ev1a even exceeded that of parental Neo enzyme by 18-fold.

The effects of fragmentation and mutagenesis are non-additive

The introduction of the R177S/V198E mutations into heterodimer H1 results in a dramatic decrease in $K_{m,kan}$. These mutations were introduced into the parental Neo enzyme, to create Neo(Ev1a), in order to address whether the mutations would have the same effect on the original full-length Neo. Contrary to expectations, these mutations had a negative effect on $K_{m,kan}$ (Table 2). Thus, the effects on Neo of the processes of (i) adding both R177S/V198E
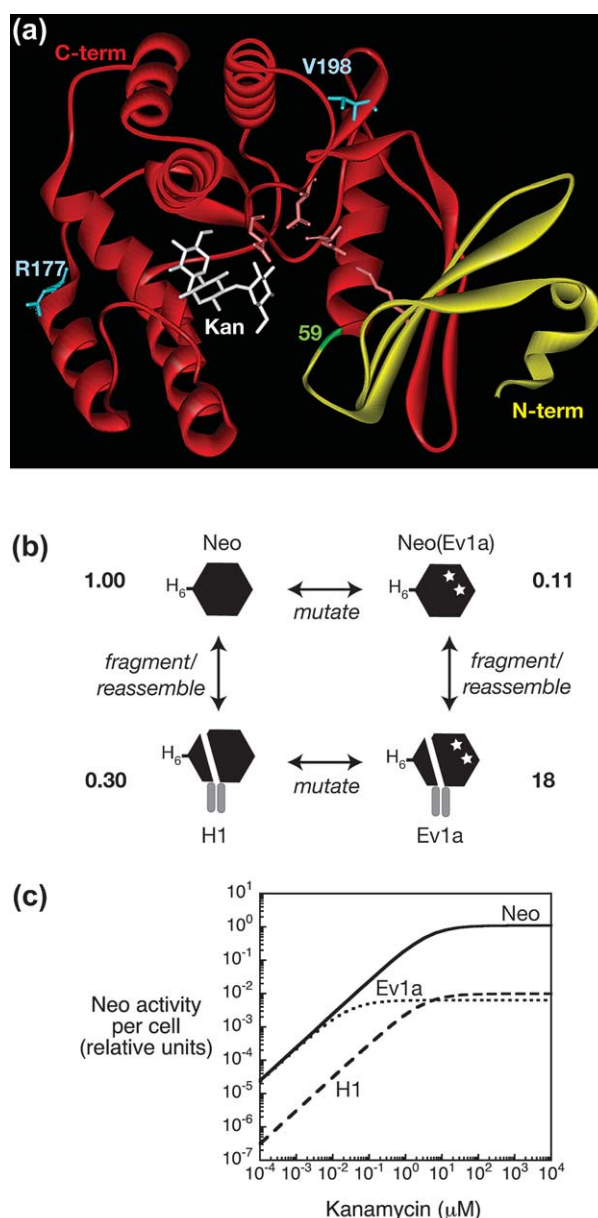

Figure 5. Characterization of Ev1a. (a) Ribbon diagram of the structure of Neo[23] bound to kanamycin (white) indicating the site of bisection for Ev1a (at residue 59; green), the N-terminal fragment (yellow) and the C-terminal fragment (red). The side-chains of R177 and V198 are shown in blue and the side-chains of key catalytic residues are shown in pink. (b) Mutagenesis cycle for creating Ev1a from Neo. Relative catalytic efficiencies ($k_{cat}/K_{m,Kan}$) normalized to that of Neo are shown. (c) Kanamycin phosphorylation velocity on a per cell basis at a saturating concentration of ATP as a function of kanamycin concentration. Velocities were calculated from the Michaelis-Menten equation using the kinetic constants from Table 2 and normalized to the production level of soluble protein. These calculations assume the absence of product. Continuous line, cells expressing Neo; long-dash broken line, cells expressing H1; short-dash broken line, cells expressing Ev1a.

mutations and (ii) fragmentation and fusion to leucine zippers are highly non-additive (Figure 5(b)). Individually, both processes have a negative effect but together their effect is positive. Additionally, "mutagenesis" of Neo(Ev1a) solely by fragmentation and fusion to leucine zippers (and not through any change in the amino acids of Neo) yields an improved enzyme with a 170-fold increase in k_{cat}/K_m. This illustrates the potential of protein fragmentation and reassembly for improving protein function. The catalytic efficiency ($k_{cat}/K_{m,kan}$) is the most relevant parameter for Neo in terms of its biological role, since activity at very low concentrations of antibiotic in the cell is believed to determine the level of antibiotic resistance (see Discussion).

Discussion

Although our future interests lie in utilizing protein fragmentation and conditional reassembly to control protein function using CIDs or other signals to control dimerization, this added level of complexity was unnecessary for the questions we sought to investigate here. The questions we sought to address were: (1) can we create heterodimers requiring dimerization domains in a systematic fashion; (2) can these non-natural heterodimers exhibit wild-type abilities *in vivo*; and (3) what is the importance of the linker between the split proteins and the domains that facilitate assembly? We chose previously characterized leucine zippers[5,22] as the domain to facilitate reassembly of fragments of Neo. We used Neo as our model protein, since it afforded an easily selectable phenotype and the fragment pairs identified might be useful in PCAs. Indeed, it has been reported that Neo can be used in a PCA in CHO cells when split near the boundary between the two structural domains (residue 99) and fused to GCN4 leucine zippers through 15 amino acid residue linkers.[4]

Systematic identification of highly active heterodimers

We used incremental truncation as a combinatorial method to locate fragments of Neo that are amenable to reassembly but only with the assistance of dimerization domains. This was aided by the fortuitous fact that there were no fragments of Neo that could appreciably assemble *in vivo* in the absence of the leucine zippers, at least when fragmented in the region between residues 51 and 207. In the libraries in which antiparallel leucine zippers were fused to fragments of Neo without a linker or with the GSGG linker, we estimate that of the order of 100 different fragment pairs in each library confer resistance to kanamycin above background levels. However, none of these fragment pairs confers the level of kanamycin resistance that wild-type Neo confers. From the perspective of controlling protein function, we desire that the

fragmented proteins retain wild-type functionality. We successfully used random mutagenesis to restore the kanamycin-resistance phenotype conferred by heterodimer H1 to wild-type levels while maintaining its dependence on leucine zippers for function *in vivo*. This demonstrated that fragmented proteins with wild-type functionalities can be constructed.

Tafelmeyer *et al*. recently demonstrated a different combinatorial method for creating protein fragment libraries involving the digestion of a circularized gene with DNaseI.[10] Their method has the advantage that it creates a library containing a higher percentage of gene fragment pairs without large deletions or overlaps between them and may be more useful for identifying active split proteins for which a strong selection is not possible. The lack of large overlap between gene fragment pairs in the library presumably also has the advantage of reducing the rate of false-positives arising by recombination between the fragments. However, the methodology is more complex and searches confined to specific regions of a protein are not feasible. Thus, large fragments that are active by themselves will be false-positives. In addition, random digestion of circular DNA molecules with DNaseI preferentially creates linear molecules with deletions at the digestion site as opposed to tandem duplications that would result in sequence overlap between the two protein fragments.[29] Sequence overlap in complementing fragments was found here and has been described in many previously described complementing protein fragments.[3,30–33] Such constructs might not be identified by the method of Tafelmeyer *et al*.[10]

Importance of minimizing linker length

Our combinatorial approach demonstrated that there are more fragment pairs of Neo that can be assembled without linkers or with short linkers than there are with longer linkers. One might have expected that the opposite would be true, that the number of fragment pairs that can assemble with longer linkers would be greater because the longer linker allows the fragments the freedom to find each other, whereas the lack of a linker would constrain the fragments in an orientation in which functional association was not possible. Undoubtedly there are some fragment pairs in which the dimerization of the leucine zippers does prevent the correct association of the fragments. However, we postulate that the function of the leucine zippers is to bring the fragments into proximity and thereby increase the local concentration of each fragment's partner (thus overcoming potential poor affinity of the two fragments for each other) as well as decreasing the conformational entropy. In addition, bringing the fragments proximal can allow the fragments to assist in each other's folding.[4] Increased linker length can be counterproductive to these mechanisms. We speculate that for many Neo fragment pairs that can assemble without the

use of linkers or with short linkers, the addition of a long linker decreases the local concentration of the two fragments relative to each other and gives the fragments too much freedom to be in geometric configurations that are not conducive to correct assembly into a functional protein. For some heterodimers that are functional with no linker or a minimal linker, we imagine that the enzyme fragments are not as free to sample a large number of geometric configurations relative to each other but rather are "presented" to each other in an orientation optimal for correct assembly. This entropic contribution should facilitate the creation of conditional heterodimers with a large difference in function between the "on" and the "off" state.

How a decrease in $K_{m,Kan}$ results in increased resistance to high concentrations of kanamycin

Characterization of both H1 and the evolved heterodimer Ev1a indicated that the major improvement was Ev1a's large decrease in K_m for kanamycin. Although the level of kanamycin on the plates (850 μM) is well above the wild-type $K_{m,Kan}$ (∼5 μM), the result is consistent with the mechanism by which cells exposed to aminoglycosides are killed. Davis has proposed that initially, only very small amounts of kanamycin are able to enter the cell, and that the ability of the enzyme to phosphorylate kanamycin when present in the cell at these low concentrations will determine cell survival.[34] This is consistent with the observations that the MIC for different aminoglycosides is correlated positively with $k_{cat}/K_{m,Kan}$ and not k_{cat}.[35,36] A cell unable to phosphorylate kanamycin efficiently as the concentration of kanamycin rises slowly from zero will produce truncated proteins in accord with kanamycin's mode of action. These truncated proteins compromise the integrity of the cell, allowing more kanamycin to enter the cell, which compromises protein synthesis even more.[34] This feed-forward mechanism accelerates toward cell-death. Higher concentrations of the aminoglycoside outside the cell result in an increased diffusion into the cell, requiring a higher level of enzyme activity to prevent cell death. As seen in Figure 5(c), Ev1a is more effective than H1 at low concentrations of kanamycin (sub-$K_{m,Kan}$). This is a direct result of the dramatic decrease in $K_{m,Kan}$, a decrease resulting from the synergistic effect of the R177S/V198E mutations and fragmentation/reassembly.

The fact that tobramycin, a competitive inhibitor that differs from kanamycin only at the 3′ hydroxyl, has a 20-fold improved K_i for Ev1a (Table 2) suggests that an increase in affinity for kanamycin is at least partially responsible for the large decrease in $K_{m,Kan}$ exhibited by Ev1a. However, the kinetic mechanism of Neo is not known and has been difficult to elucidate.[24] Thus, it is possible that the reasons for a decreased $K_{m,Kan}$ involve steps in the reaction pathway other than substrate binding. Furthermore, the products of the reaction will be present *in vivo* and the MIC could be a function of the extent to which the enzymes exhibit product inhibition. However, regardless of the mechanism, Ev1a is superior to H1 *in vivo* and Ev1a's $k_{cat}/K_{m,Kan}$ is superior to that of Neo, H1 or Neo(Ev1a).

The role of fragmentation and reassembly in improving function

The wild-type sequence of Neo is catalytically superior when comparing the non-fragmented proteins (Neo and Neo(Ev1a)). The mutated R177S/V198E sequence is catalytically superior when comparing the fragmented proteins (H1 and Ev1a). The only difference between the fragmented and non-fragmented versions is a break in the peptide backbone that is replaced by the anti-parallel leucine zipper pairs and the duplication of residue 59. Presumably, the effect of fragmentation and reassembly manifests itself in a slightly altered structure of the Neo domain, such that mutations that would otherwise be deleterious to function in Neo are now advantageous in H1. Increased flexibility afforded by the break of the peptide bond could play a role as well. In the absence of structural information on the heterodimers, it is difficult to speculate on the reasons for the different effects of the mutations on the heterodimer and the parental protein. An important related point is that the differences between Neo(Ev1a) and Ev1a are not in the sequence of the Neo domain (aside from the duplication of residue 59) but rather in the fragmentation of this sequence and the addition of new sequences (the leucine zippers) to the Neo domain. Fragmentation and the addition of sequences causes a large jump to a new but related area of sequence space, one in which the relationship between the sequence in the Neo domain and its function ($k_{cat}/K_{m,Kan}$) has been altered.

Fragmentation and reassembly as a potential strategy for improving function

To our knowledge, this is the first example of improving protein function through fragmentation. Like mutagenesis (e.g. point mutagenesis), fragmentation will often result in compromised or abolished activity through structural changes and disruption of other phenomena important for protein function (e.g. protein folding, stability and dynamics). However, there is no reason, *a priori*, that fragmentation necessarily results in compromised activity. When viewed as a form of mutagenesis, it seems plausible that fragmentation might, in some cases, result in improved activity. Whether the approach has any general applicability remains to be seen. Ev1a's 170-fold and 18-fold increase in catalytic efficiency over Neo(Ev1a) and wild-type Neo, respectively, indicate that the approach has the potential of being very effective.

Materials and Methods

Plasmid construction

Plasmids pDIM-N2 and pDIM-C8 are compatible plasmids that confer ampicillin resistance and chloramphenicol resistance, respectively.[19] The *neo* gene was amplified from pSV2-Neo using overlap extension PCR (such that an internal NcoI restriction site was removed) and cloned between the NdeI and SpeI sites of pDIM-N2 downstream from the IPTG-inducible *lac* promoter to create pDIM-N2-Neo. An N-terminal fragment of *neo* that coded for amino acid residues 1–207 (Neo[1-207]) was PCR amplified from pDIM-N2-Neo and cloned between the NdeI and BamHI sites of pDIM-N2 to create pDIM-N2-Neo[1-207]. Similarly, a C-terminal fragment of *neo* that coded for amino acid residues 51–264 (Neo[51-264]) was PCR amplified from pDIM-N2-Neo and cloned between the BglII and SpeI sites of pDIM-C8 to create pDIM-C8-Neo[51-264].

DNA coding for antiparallel leucine zippers NZ and CZ[5] was PCR amplified and cloned into pDIM-N2-Neo[1-207] and pDIM-C8-Neo[51-264], respectively. The DNA for NZ was inserted between the NsiI and KpnI sites such that incremental truncation would result in fusion of the following sequence to the C terminus of C-terminal truncations of Neo[1-207]: SAQLKKEL QANKKEL AQLKWELQALKKELA Q. GSGG or (GSGG)$_2$ linkers, if included, were inserted after the first serine residue. The DNA for CZ was inserted between the NcoI and SacI sites such that incremental truncation would result in fusion of the following sequence to the N terminus of N-terminal truncations of Neo[51-264]: MASAQLEKKLQALEK KLAQLEWKNQALEKK LAQ. GSGG or (GSGG)$_2$ linkers, if included, were appended to the C terminus of this sequence. Analogous constructs with sequences coding for parallel leucine zippers P-NZ (SAQLEKELQALEKEN AQLEWELQALEKELA Q) and P-CZ (MASAQLKKKLQALKK KNAQLKWKLQALKKK LAQ) with linker sequences coding for (GSGG)$_3$ linkers were also constructed.

Preparation of heterodimer libraries

Incremental truncation libraries were created as described[19] and as depicted in Figure 1. Following the creation of N-terminal and C-terminal incremental truncation libraries, 20 ng of each library was cotransformed into DH5α-E cells and plated onto a large LB agar plate (245 mm × 245 mm) supplemented with ampicillin and chloramphenicol.

ELISA analysis of soluble protein fractions

ELISA analysis was performed on the soluble fraction of cell lysates using anti-neomycin phosphotransferase type II polyclonal antibody (Cortex Biochem, San Leandro, CA) as the primary antibody and goat anti-rabbit IgG-alkaline phosphatase conjugate (Bio-Rad, Hercules, CA) as the secondary antibody. Cells were grown at 37 °C in 50 ml of LB medium overnight, followed by centrifugation at 5000 rpm. The supernatant was discarded and the cell pellet was resuspended in 5 ml of cell lysis buffer I (50 mM Tris–HCl (pH 8.0), 1 mM EDTA (pH 8.0), 100 mM NaCl). The cells were lysed with one pass through a French pressure cell at 20,000 psi (1 psi ∼6.9 kPa). The lysate was centrifuged at ≥20,000*g* to separate the soluble from the insoluble fraction, with the soluble fraction in the supernatant. The total protein concentration of the soluble fraction was determined using the *DC* Protein Assay Kit (Bio-Rad, Hercules, CA). Dilutions of the soluble fraction (100 µl containing 1.75 µg of total soluble protein) were added to a 96-well plate and incubated for 1 h. The wells were emptied and incubated for 5 min with 250 µl of blocking solution containing 2% (w/v) gelatin in TBST (10 mM Tris–HCl (pH 7.5), 150 mM NaCl, 0.005% (v/v) Tween 20). The wells were emptied and incubated with 100 µl of anti-Neo antibody (diluted 1:1000 (v/v) in blocking solution) for 1 h. The wells were then washed five times with TBST and incubated for 1 h with 100 µl of secondary antibody (diluted 1:3000 (v/v) in blocking solution). The wells were washed five times with TBST. A total of 100 µl of developing solution (10 mg of *p*-nitrophenyl phosphate in 1.02 M diethanolamine) was added to each well and the plate was incubated for 30 min at room temperature. The absorbance at 405 nm was then measured. All samples were analyzed in triplicate from three independent cultures. A standard curve was constructed by mixing in different ratios of soluble protein from cells expressing wild-type Neo and cells not containing any *neo* gene or gene fragment. Thus, all production levels are expressed relative to that of wild-type Neo.

Purification of wild-type Neo and heterodimers

DNA coding for the amino acid sequence GSSH$_6$-SSGLVPRGSH was added by PCR to the beginning of any gene whose protein was purified. All enzymes were purified using the His·Bind Purification Kit (Novagen, Madison, WI). Wild-type Neo was purified from a 1 l culture grown to an OD$_{600\,nm}$ of 0.5 followed by induction for 1 h with 0.1 mM IPTG. H1 and Ev1a were purified from cells grown for 8 h in a 5 l BBI Biostat B fermenter and induced for 1 h with 1.0 mM IPTG prior to purification. Cells were lysed either by French press or by sonication and the soluble fraction of the lysate was loaded onto the purification column. After elution with imidazole, the proteins were dialyzed at 4 °C against 1 l of 20 mM Hepes, 20 mM β-mercaptoethanol, 10% (v/v) glycerol for 4 h (twice) followed by dialysis against 1 l of the same solution overnight. The purified protein was stored as described.[24]

Determination of kinetic constants of purified enzymes

Kinetic constants were determined at 25 °C using a spectrofluorometric assay that couples aminoglycoside phosphorylation to NADH oxidation,[25] using the slight modifications described by Siregar *et al.*[24] Assays were performed using 10–60 nM protein on a QuantaMax 4 spectrofluorometer (Photon Technologies, Lawrenceville, NJ). The range of substrate concentrations was 30–80 µM for ATP and 1.5–120 µM for kanamycin. Initial rates from a minimum of 12 assays and six different substrate concentrations were used for each determination of k_{cat} and K_m. The kinetic parameters and errors were determined using Lenora version 1.0.[37] Errors for k_{cat} include an estimated 10% error in protein concentration. The kinetic properties of heterodimer Ev1a were

determined with the aid of the competitive inhibitor tobramycin.

Acknowledgements

References

1. Kato, I. & Anfinsen, C. B. (1969). On the stabilization of ribonuclease S-protein by ribonuclease S-peptide. *J. Biol. Chem.* **244**, 1004–1007.

2. Ullmann, A., Jacob, F. & Monod, J. (1967). Characterization by *in vitro* complementation of a peptide corresponding to an operator-proximal segment of the beta-galactosidase structural gene of *Escherichia coli*. *J. Mol. Biol.* **24**, 339–343.

3. Ostermeier, M., Nixon, A. E., Shim, J. H. & Benkovic, S. J. (1999). Combinatorial protein engineering by incremental truncation. *Proc. Natl Acad. Sci. USA*, **96**, 3562–3567.

4. Michnick, S. W., Remy, I., Campbell-Valois, F. X., Vallee-Belisle, A. & Pelletier, J. N. (2000). Detection of protein–protein interactions by protein fragment complementation strategies. *Methods Enzymol.* **328**, 208–230.

5. Ghosh, I., Hamilton, A. D. & Regan, L. (2000). Antiparallel leucine zipper-directed protein reassembly: application to the green fluorescent protein. *J. Am. Chem. Soc.* **122**, 56–58.

6. Pelletier, J. N., Campbell-Valois, F. X. & Michnick, S. W. (1998). Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments. *Proc. Natl Acad. Sci. USA*, **95**, 12141–12146.

7. Rossi, F., Charlton, C. A. & Blau, H. M. (1997). Monitoring protein–protein interactions in intact eukaryotic cells by β-galactosidase complementation. *Proc. Natl Acad. Sci. USA*, **94**, 8405–8410.

8. Karimova, G., Pidoux, J., Ullmann, A. & Ladant, D. (1998). A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc. Natl Acad. Sci. USA*, **95**, 5752–5756.

9. Paulmurugan, R., Umezawa, Y. & Gambhir, S. S. (2002). Noninvasive imaging of protein-protein interactions in living subjects by using reporter protein complementation and reconstitution strategies. *Proc. Natl Acad. Sci. USA*, **99**, 15608–15613.

10. Tafelmeyer, P., Johnsson, N. & Johnsson, K. (2004). Transforming a (beta/alpha)8–barrel enzyme into a split-protein sensor through directed evolution. *Chem. Biol.* **11**, 681–689.

11. Johnson, N. & Varshavsky, A. (1994). Split ubiquitin as a sensor of protein interactions *in vivo*. *Proc. Natl Acad. Sci. USA*, **91**, 10340–10344.

12. Galarneau, A., Primeau, M., Trudeau, L. E. & Michnick, S. W. (2002). Beta-lactamase protein fragment complementation assays as *in vivo* and *in vitro* sensors of protein protein interactions. *Nature Biotechnol.* **20**, 619–622.

13. Wehrman, T., Kleaveland, B., Her, J. H., Balint, R. F. & Blau, H. M. (2002). Protein–protein interactions monitored in mammalian cells *via* complementation of beta-lactamase enzyme fragments. *Proc. Natl Acad. Sci. USA*, **99**, 3469–3474.

14. Remy, I., Wilson, I. A. & Michnick, S. W. (1999). Erythropoietin receptor activation by a ligand-induced conformation change. *Science*, **283**, 990–993.

15. Remy, I. & Michnick, S. W. (1999). Clonal selection and *in vivo* quantitation of protein interactions with protein-fragment complementation assays [published erratum appears in *Proc. Natl Acad. Sci. USA*, 1999 Jun 22;96(13):7610]. *Proc. Natl Acad. Sci. USA*, **96**, 5394–5399.

16. Mootz, H. D., Blum, E. S., Tyszkiewicz, A. B. & Muir, T. W. (2003). Conditional protein splicing: a new tool to control protein structure and function *in vitro* and *in vivo*. *J. Am. Chem. Soc.* **125**, 10561–10569.

17. Mootz, H. D. & Muir, T. W. (2002). Protein splicing triggered by a small molecule. *J. Am. Chem. Soc.* **124**, 9044–9045.

18. Spencer, D. M., Wandless, T. J., Schreiber, S. L. & Crabtree, G. R. (1993). Controlling signal transduction with synthetic ligands. *Science*, **262**, 1019–1024.

19. Paschon, D. & Ostermeier, M. (2004). Construction of protein fragment complementation libraries using incremental truncation. *Methods Enzymol.* **388**, 103–116.

20. Hon, W. C., McKay, G. A., Thompson, P. R., Sweet, R. M., Yang, D. S., Wright, G. D. & Berghuis, A. M. (1997). Structure of an enzyme required for amino-glycoside antibiotic resistance reveals homology to eukaryotic protein kinases. *Cell*, **89**, 887–895.

21. Boehr, D. D., Thompson, P. R. & Wright, G. D. (2001). Molecular mechanism of aminoglycoside antibiotic kinase APH(3′)-IIIa: roles of conserved active site residues. *J. Biol. Chem.* **276**, 23929–23936.

22. O'Shea, E. K., Lumb, K. J. & Kim, P. S. (1993). Peptide 'Velcro': design of a heterodimeric coiled coil. *Curr. Biol.* **3**, 658–667.

23. Nurizzo, D., Shewry, S. C., Perlin, M. H., Brown, S. A., Dholakia, J. N., Fuchs, R. L. *et al.* (2003). The crystal structure of aminoglycoside-3′-phosphotransferase-IIa, an enzyme responsible for antibiotic resistance. *J. Mol. Biol.* **327**, 491–506.

24. Siregar, J. J., Lerner, S. A. & Mobashery, S. (1994). Purification and characterization of aminoglycoside 3′- phosphotransferase type IIa and kinetic comparison with a new mutant enzyme. *Antimicrob. Agents Chemother.* **38**, 641–647.

25. Perlin, M. H., McCarty, S. C. & Greer, J. P. (1988). Coupled spectrofluorometric assay for aminoglycoside phosphotransferases. *Anal. Biochem.* **171**, 145–149.

26. Magliery, T. J., Wilson, C. G., Pan, W., Mishler, D., Ghosh, I., Hamilton, A. D. & Regan, L. (2005). Detecting protein–protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism. *J. Am. Chem. Soc.* **127**, 146–157.

27. Caldwell, R. C. & Joyce, G. F. (1995). Mutageneic PCR.

In *PCR Primer: A Laboratory Manual* (Dieffenbach, C., ed.), pp. 583–589, Cold Spring Harbor Laboratory Press, Plainview, NY.

28. Shafikhani, S., Siegel, R. A., Ferrari, E. & Schellenberger, V. (1997). Generation of large libraries of random mutants in *Bacillus subtilis* by PCR-based plasmid multimerization. *Biotechniques*, **23**, 304–310.

29. Guntas, G. & Ostermeier, M. (2004). Creation of an allosteric enzyme by domain insertion. *J. Mol. Biol.* **336**, 263–273.

30. Hantgan, R. R. & Taniuchi, H. (1977). Formation of a biologically active, ordered complex from two overlapping fragments of cytochrome c. *J. Biol. Chem.* **252**, 1367–1374.

31. Matsuyama, S., Kimura, E. & Mizushima, S. (1990). Complementation of two overlapping fragments of SecA, a protein translocation ATPase of *Escherichia coli*, allows ATP binding to its amino-terminal region. *J. Biol. Chem.* **265**, 8760–8765.

32. Posfai, G., Kim, S. C., Szilak, L., Kovacs, A. & Venetianer, P. (1991). Complementation by detached parts of GGCC-specific DNA methyltransferases. *Nucl. Acids Res.* **19**, 4843–4847.

33. Taniuchi, H. & Anfinsen, C. B. (1971). Simultaneous formation of two alternative enzymology active structures by complementation of two overlapping fragments of staphylococcal nuclease. *J. Biol. Chem.* **246**, 2291–2301.

34. Davis, B. D. (1987). Mechanism of bactericidal action of aminoglycosides. *Microbiol. Rev.* **51**, 341–350.

35. McKay, G. A., Thompson, P. R. & Wright, G. D. (1994). Broad spectrum aminoglycoside phosphotransferase type III from Enterococcus: overexpression, purification, and substrate specificity. *Biochemistry*, **33**, 6936–6944.

36. McKay, G. A. & Wright, G. D. (1995). Kinetic mechanism of aminoglycoside phosphotransferase type IIIa. Evidence for a Theorell-Chance mechanism. *J. Biol. Chem.* **270**, 24686–24692.

37. Cornish-Bowden, A. (1995). *Analysis of Enzyme Kinetic Data*, Oxford University Press, New York.