(27) or a recombinant c-*myc*– or *lacZ*-expressing adenovirus. We constructed c-*myc*–expressing adenovirus by cloning the full-length human c-*myc* cDNA into a replication-deficient Ad5 as described for *lacZ*-expressing adenovirus (28).
19. H. E. Ruley, *Nature* **304**, 602 (1983); S. W. Lowe, H. E. Ruley, T. Jacks, D. E. Housman, *Cell* **74**, 957 (1993); M. E. McCurrach, T. M. Connor, C. M. Knudson, S. J. Korsmeyer, S. W. Lowe, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 2345 (1997).
20. Generation of E1A- and Ras-expressing MEFs was accomplished by mass infection of early-passage MEFs with E1A 13S- and H-rasV12–expressing retroviral vectors as described (19).
21. M. Adachi *et al.*, *Nature Genet.* **11**, 294 (1995); K. Pfeffer *et al.*, *Cell* **73**, 457 (1993).
22. C. Moens, B. R. Stanton, L. F. Parada, J. Rossant, *Development* **119**, 485 (1993).
23. T. P. Yamaguchi, D. J. Dumont, R. A. Conlon, M. L. Breitman, J. Rossant, *Dev. Suppl.* **118**, 489 (1993).
24. Y. Mishina, A. Suzuki, N. Ueno, R. R. Behringer, *Genes Dev.* **9**, 3027 (1995).
25. R. M. Mortensen, D. A. Conner, S. Chao, A. A. Geisterfer-Lowrance, J. G. Seidman, *Mol. Cell. Biol.* **12**, 2391 (1992).
26. I. Schmid, C. H. Uittenbogaart, J. V. Giorgi, *Cytometry* **15**, 12 (1994).
27. E. White and B. Stillman, *J. Virol.* **61**, 426 (1987).
28. W. S. El-Deiry *et al.*, *Cell* **75**, 817 (1993).
29. M. E. McCurrach and S. W. Lowe, unpublished data.
30. We thank V. M. Dixit for DR3 and DR4 vectors and P. H. Krammer for agonistic antibody to Fas; H. Hsu, L.-M. Boucher, and members of T.W.M.'s laboratory for material and instructive discussions; I. del Barco Barrantes and B. Hum for technical assistance; M. Saunders for scientific editing; and I. Ng for administrative support. Supported in part by a grant from the National Cancer Institute of Canada (NCIC) and grant CA13106 from the National Cancer Institute and Medical Research Council of Canada.

14 January 1998; accepted 18 February 1998

Redesigning Enzyme Topology by Directed Evolution

Gavin MacBeath, Peter Kast, Donald Hilvert*

Genetic selection was exploited in combination with structure-based design to transform an intimately entwined, dimeric chorismate mutase into a monomeric, four-helix–bundle protein with near native activity. Successful reengineering depended on choosing a thermostable starting protein, introducing point mutations that preferentially destabilize the wild-type dimer, and using directed evolution to optimize an inserted interhelical turn. Contrary to expectations based on studies of other four-helix–bundle proteins, only a small fraction of possible turn sequences (fewer than 0.05 percent) yielded well-behaved, monomeric, and highly active enzymes. Selection for catalytic function thus provides an efficient yet stringent method for rapidly assessing correctly folded polypeptides and may prove generally useful for protein design.

Several recent studies have focused on reducing the size of protein domains to obtain minimized functional units either by removing secondary structural elements within a domain (*1*) or by converting oligomeric proteins into monomers (*2*, *3*). The latter is particularly challenging when the individual polypeptides are intricately entwined. Monomeric variants of such dimers have been constructed for interleukin-5 by inserting a loop sequence from a homologous protein (*2*) and for λ cro by screening a small subset of carefully designed variants (*3*). A high-resolution crystal structure of the monomeric cro protein, however, revealed significant conformational changes relative to the wild-type dimer (*4*).

The sensitivity of catalytic efficiency to structural perturbations makes minimization of oligomeric enzymes especially difficult. For example, although the active sites of the homodimeric enzyme triosephosphate isomerase do not contain residues from both polypeptides, disruption of the dimer interface results in a >10,000-fold reduction in k_{cat}/K_m (where k_{cat} and K_m are the catalytic rate and Michaelis constants, respectively) (*5*, *6*). Enzymes with active sites at the subunit interface pose even greater challenges. In this context, we considered chorismate mutase (CM), which catalyzes the conversion of chorismate to prephenate in the biosynthesis of L-tyrosine (Tyr) and L-phenylalanine (Phe) with rate accelerations of about 10^6 (*7*). Members of the AroQ class (*8*) of CM are found in a wide variety of bacteria, fungi, and higher plants. Crystallographic studies on a typical AroQ enzyme from *Escherichia coli* (EcCM) have revealed an intriguing homodimeric topology, in which each catalytic domain adopts a four-helix–bundle structure (*9*) (Fig. 1). The long amino-terminal helix (H1) spans the two domains and contributes residues to both active sites of the dimer. We reasoned that if the appropriate loop were inserted in the middle of this helix, a 180° turn could be induced that would allow the amino-terminal half of H1 to fold back and displace the other polypeptide in the dimer to form a monomer (Fig. 1).

The choice of loop sequence is of considerable interest. Previous studies on other four-helix–bundle proteins have shown that there are surprisingly few constraints on the sequence or length of interhelical turns (*10*–*12*). This finding suggests that almost any turn could be used to convert the CM domain into a functional monomer. However, we have recently shown that if residues in an interhelical turn are involved in long-range tertiary interactions, the fraction of acceptable turn sequences can be dramatically reduced (*13*). In the ideal monomer, the inserted loop not only must adopt the appropriate conformation, it also must interact favorably with the newly exposed hydrophobic regions that result from disruption of the subunit interface. Given the exacting conformational control required for efficient enzyme catalysis, prediction of an optimal turn sequence becomes intractable.

By introducing randomized sequences into the AroQ gene and transforming CM-deficient bacteria with the resulting library, acceptable turn sequences that confer high activity should be selectable by genetic complementation (*13*). When this approach was used with EcCM, however, we were unable to obtain stable, monomeric variants (*14*). Although the enzyme active site is preserved in this topological transformation, extensive hydrophobic interactions formed at the subunit interface are lost. This interface comprises much of the core of the folded dimer and hence should contribute significantly to the overall stability of the protein. In support of this notion, denaturation studies show that EcCM undergoes a cooperative transition from folded dimer to unfolded monomer without accumulation of a structured, monomeric intermediate. Analogous results have been obtained with other dimeric proteins in which quaternary structure accounts for most, if not all, of the overall stability (*15*).

We therefore sought a more stable version of the AroQ CM domain. Recent progress in whole genome sequencing has provided access to a plethora of uncharacterized open reading frames (ORFs) from a wide variety of organisms. The complete genomic sequence of the thermophilic archaeon *Methanococcus jannaschii* (*16*) revealed an ORF for a protein with high similarity to EcCM (Fig. 2). No insertions or deletions are apparent, and the active site residues are remarkably conserved. We have subcloned and overexpressed this ORF

The Scripps Research Institute, Departments of Chemistry and Molecular Biology, 10550 North Torrey Pines Road, La Jolla, California, 92037, USA.

*To whom correspondence should be addressed at Department of Chemistry, Swiss Federal Institute of Technology (ETH), CH-8092 Zürich, Switzerland. E-mail: hilvert@org.chem.ethz.ch

(which we now refer to as the $aroQ_f$ gene) and have characterized the resulting protein, MjCM. Its circular dichroism spectrum matches that of EcCM, it exists as a homodimer, and it promotes the rearrangement of chorismate to prephenate with high efficiency ($k_{cat}/k_{uncat} = 1.1 \times 10^6$ at 20°C) (*17*). In addition, the MjCM- and EcCM-catalyzed reactions exhibit comparable energy of activation parameters (*15*), further demonstrating the similarity between these two proteins. They differ, however, in their intrinsic stabilities. MjCM exhibits a midpoint for thermal denaturation (T_m) that is 25°C higher than that for EcCM (at 1 μM concentration), and chemical denaturation studies indicate that MjCM is about 5 kcal/mol more stable than EcCM (*15*). Because the subunit interface of the two proteins is highly conserved, much of this additional stability must reside within elements of the structure not directly associated with dimerization.

We based our experimental design on a homology model of MjCM. Structural comparison with a helix-turn-helix motif in seryl-tRNA synthetase (*18*, *19*) suggested duplicating two residues in helix H1 (Leu^{20}-Lys^{21}) and introducing a six-residue turn between the repeated units (that is, -Leu^{20}-Lys^{21}-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Leu^{20a}-Lys^{21a}-). Two point mutations (Leu^{20a} to Glu and Ile^{77} to Arg) were also introduced to destabilize the dimer without affecting the resulting monomer (Fig. 1). Working at the genetic level, we cloned the gene for MjCM into a high copy number expression vector, made the aforementioned modifications, and prepared a combinatorial library of variants by inserting six random codons at the appropriate position (*20*). The pool of plasmids was first introduced into prototrophic *E. coli* cells (>10^8 transformants) and amplified under nonselective conditions. Expression of the recombinant gene is repressed in this strain of *E. coli*, so all the transformants should replicate at approximately the same rate. The DNA recovered from these cells was termed "unselected."

The unselected DNA was then introduced into *E. coli* strain KA12/pKIMP-UAUC, which lacks both endogenous CM genes (*21*). Without CM activity, KA12/pKIMP-UAUC cells are unable to synthesize Tyr and Phe and hence cannot grow on minimal medium (M9c) that lacks these essential amino acids. When a plasmid encoding a functional CM is introduced into these cells, they regain their prototrophy (*21*). The transformed cells were initially amplified about 100-fold in rich medium (LB broth) to select against toxic variants. We previously noticed that KA12/pKIMP-UAUC cells producing the wild-type MjCM protein grow slowly, even on rich medium. We ascribe this toxicity to problems with protein aggregation, because overproduction of MjCM using T7-based vectors yields about 25% folded dimer and 75% misfolded, higher order aggregates. After this first round of selection for proteins not interfering with cellular function, the collection of transformants was amplified about 500-fold in minimal medium (M9c) lacking both Tyr and Phe. By selecting for CM activity in liquid culture, all the members of the library are in direct competition with each other and the fastest growing clones are amplified to a greater extent than the "weaker" variants. The DNA recovered from this evolution experiment was termed "selected."

The effectiveness of this selection experiment was assessed by transforming fresh KA12/pKIMP-UAUC cells with either the unselected or the selected DNA. Transformed cells were plated both on rich medium (LB) and on minimal medium (M9c) (Fig. 3A). After 3 days at 30°C, only 0.7% of the clones from the unselected DNA produced large colonies (>0.75 mm diameter) on M9c plates. In contrast, 81% of the clones from the selected DNA yielded large colonies. This represents an enrichment of >100-fold in variants exhibiting high CM activity.

To identify monomeric variants of MjCM, we transferred the pool of selected genes to the T7-based expression vector pET-22b(+) (Novagen) and isolated individual clones for further analysis (*22*). His_6-tagged proteins were purified by affinity chromatography and their quaternary structures were analyzed by size-exclusion column chromatography. Unlike wild-type MjCM, none of the 26 proteins that were screened showed evidence of higher order aggregates. However, most of the proteins appeared to be mixtures of monomeric and dimeric species. This situation resembles that of the dual-mode quaternary structure of bovine seminal ribonuclease (BS-RNase), which can be separated into rapidly equilibrating monomers and dimers and structurally distinct, domain-swapped dimers that interconvert at a very slow rate (half-life of hours to days at 37°C) (*23*).

In contrast to this pattern, one of the proteins, which we termed mMjCM, eluted

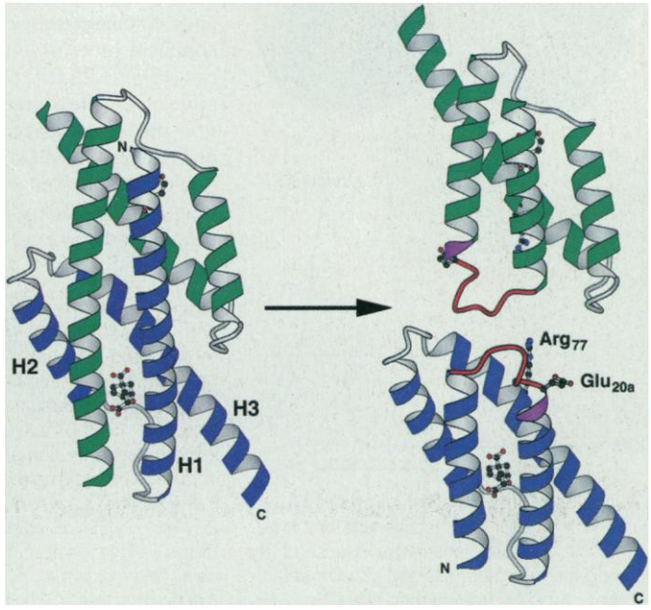

Fig. 1. Redesigning enzyme topology. The wild-type, dimeric AroQ protein is shown on the left, with transition state analog inhibitor bound in the active sites (*9*). Design of a monomeric version included duplication of two residues (pink), introduction of two point mutations (Leu^{20a} to Glu, Ile^{77} to Arg), and insertion of a randomized six-residue turn (red).

```
MTSENPLLALREKISALDEKLLALLAERRELAVEVGKAKLLSHRPVRDID  50
    : |  :| ||  :|:|:| |:|||  || :|:  |      |: | :
..MIEKLAEIRKKIDEIDNKILKLIAERNSLAKDVAEIKNQLGIPINDPE  48

RERDLLERLITLGKAHHLDAHYITRLFQLIIEDSVLTQQALLQQHL..... 96
||: : :|:  | | |::| :   ::||::||     |    |::
REKYIYDRIRKLCKEHNVDENIGIKIFQILIEHNKALQKQYLEETQNKNKK 99
```

Fig. 2. Sequence alignment of EcCM (top strand, black) and MjCM (bottom strand, blue). Identical and similar residues are indicated by vertical lines and colons, respectively. Residues that line the active site, as identified in the crystal structure of EcCM complexed with a transition state analog (*9*), are shown in red.

from the size-exclusion column as a single peak with the retention time expected for a monomeric species (Fig. 3B). The combination of our selection results (>100-fold enrichment) and the quaternary structure analysis (1 in 26) shows that <0.05% of turn sequences yield well-behaved, monomeric proteins. This contradicts the simple expectation that most interhelical turn sequences are functionally equivalent (*10–12*). DNA sequencing identified the six-residue insert as Ala-Arg-Trp-Pro-Trp-Ala. This fairly apolar sequence suggests that the turn may well interact with the exposed hydrophobic face created by disruption of the subunit interface. The presence of a proline residue is also compatible with a turn structure and may be important in breaking the long amino-terminal helix.

The results of analytical ultracentrifugation on mMjCM (*24*) fit well to a single ideal species model with an average molecular mass consistent with a monomeric species (Fig. 4A). At high concentrations, the data fit best to an association model involving a monomer-dimer equilibrium (K_a = 5×10^4 M^{-1}). Interconversion between quaternary forms is extremely rapid because the protein elutes as a single, smooth peak from a size-exclusion column (with shorter retention times observed at higher protein concentrations). In contrast, proteins that are known to interconvert between monomeric and domain-swapped dimeric forms do so very slowly (typically over days) and are easily separable by gel filtration (*23, 25–27*). This observation suggests that dimerization of mMjCM does not involve domain swapping (as in Fig. 1) but instead occurs by simple association of two monomers, mediated by the exposed hydrophobic face of the protein.

Like MjCM, mMjCM exhibits a far-ultraviolet circular dichroism (CD) spectrum that does not change with protein concentration and is characteristic of a predominantly α-helical protein. The strong signal at 222 nm ($[\theta] = -24{,}000$ deg cm^2 dmol^{-1}), indicative of α-helical structure, was used to monitor the urea-induced denaturation of the protein (*28*). mMjCM undergoes a cooperative transition from folded to unfolded states that is independent of protein concentration. The data fit well to a two-state model of unfolding, yielding a value of 2.7 ± 0.1 kcal/mol for the free energy of unfolding $\Delta G_U(H_2O)$. This value is fairly low compared with typical proteins (6 to 14 kcal/mol) (*15*), consistent with the notion that wild-type CM derives most of its stability from dimerization.

When assayed for catalytic activity (*29*), mMjCM exhibited saturation kinetics with a k_{cat} of 3.2 s^{-1} and a K_m of 170 μM at 20°C. These values compare remarkably well with those for wild-type MjCM (k_{cat} = 3.2 s^{-1}, K_m = 50 μM) and for EcCM (k_{cat} = 9.0 s^{-1}, K_m = 300 μM) (*17*). We transferred the gene for mMjCM back into the expression vector used for the selection experiments, and its ability to complement the CM auxotrophy of KA12/pKIMP-UAUC cells was compared with that of wild-type *aroQ*$_f$ (Fig. 4B). Although cells producing MjCM grew slowly on both nonselective and selective plates, cells producing mMjCM grew unhindered under both conditions. The selective pressure was therefore effective in both minimizing toxicity and maximizing catalytic efficiency.

Proteins with redesigned topology may serve as model systems for studying how structure, stability, and function interrelate. It is interesting to note that all natural CMs studied to date are multimeric, and investigations of the monomeric variant may also contribute to an understanding of the evolution of multimeric proteins (*25, 30*). More generally, the success of this study supports broader application of selection methods for reengineering existing enzymes and designing proteins de novo.

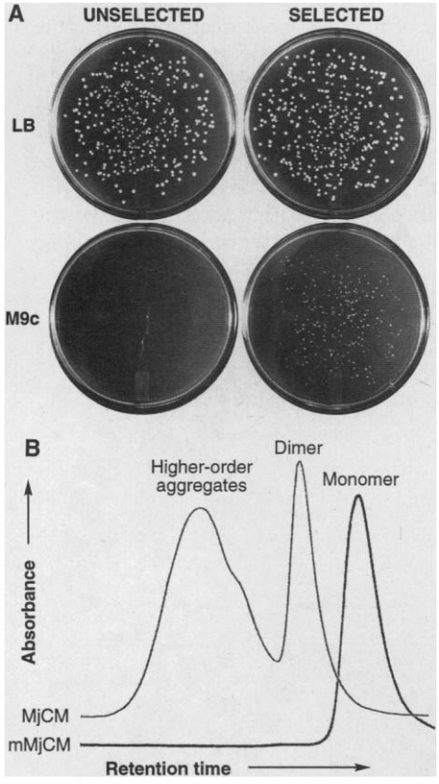

Fig. 3. Selection and identification of a highly active, monomeric CM. (**A**) Effectiveness of the genetic selection experiments was assessed by comparing the unselected library with the selected library. Freshly transformed KA12/pKIMP-UAUC cells were washed and plated in duplicate on rich medium (LB) or minimal medium (M9c). LB plates are shown after 3 days at room temperature; M9c plates are shown after 3 days at 30°C. (**B**) Monomeric variant mMjCM was identified from the selected population by analytical size-exclusion chromatography on a Superose 12 (10/30) fast protein liquid chromatography column. Overproduced protein was purified in one step by affinity chromatography and injected onto the column. Wild-type MjCM was found to be a mixture of dimer and higher order aggregates, whereas mMjCM was exclusively monomeric. Calibration of the column with known protein standards enabled assignment of the peaks.

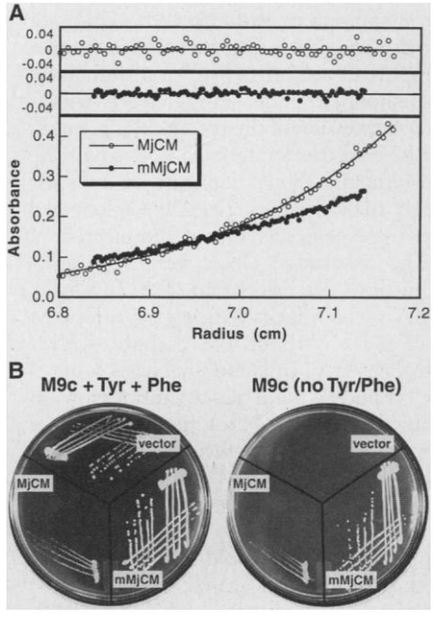

Fig. 4. Characterization of mMjCM and comparison with MjCM (*32*). (**A**) Aggregation states of mMjCM and MjCM were determined by analytical ultracentrifugation. Lower plot shows sedimentation equilibrium data (absorbance at 225 nm) collected at 17,000 revolutions per minute with a loading concentration of 2 μM protein. Both data sets were fit to a single ideal species model using the equation $A_r = \exp[\ln(A_0) + (M(1 - \bar{\nu}\rho)\omega^2/2RT)\cdot(x^2 - x_0^2)] + E$, where A_r is the absorbance at radius x, A_0 is the absorbance at a reference radius x_0 (the meniscus), M is the molecular mass of the single species, $\bar{\nu}$ is the partial specific volume of the protein, ρ is the density of the solvent, ω is the angular velocity of the rotor, R is the gas constant, T is the absolute temperature, and E is a baseline error correction factor. The fit for mMjCM gave an average molecular mass of 14,700 ± 850 daltons (13,174 daltons expected for the monomer), and the fit for MjCM gave 23,000 ± 1100 daltons (24,212 daltons expected for the dimer). The residuals for both fits are shown in the upper windows. (**B**) The ability of mMjCM and MjCM to complement the CM deficiency of KA12/pKIMP-UAUC cells was assessed by streaking the clones on minimal medium (M9c) either supplemented with Tyr and Phe (left) or lacking Tyr and Phe (right). The vector lacking a CM gene was included as a negative control.

REFERENCES AND NOTES

1. A. C. Braisted and J. A. Wells, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5688 (1996).
2. R. R. Dickason and D. P. Huston, *Nature* **379**, 652 (1996).
3. M. C. Mossing and R. T. Sauer, *Science* **250**, 1712 (1990).
4. R. A. Albright, M. C. Mossing, B. W. Matthews, *Biochemistry* **35**, 735 (1996).
5. T. V. Borchert, R. Abagyan, R. Jaenicke, R. K. Wierenga, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1515 (1994).
6. W. Schliebs, N. Thanki, R. Jaenicke, R. K. Wierenga, *Biochemistry* **36**, 9655 (1997).
7. E. Haslam, *Shikimic Acid: Metabolism and Metabolites* (Wiley, New York, 1993).
8. W. Gu *et al.*, *Microb. Comp. Genomics* **2**, 141 (1997).
9. A. Y. Lee, P. A. Karplus, B. Ganem, J. Clardy, *J. Am. Chem. Soc.* **117**, 3627 (1995).
10. A. P. Brunet *et al.*, *Nature* **364**, 355 (1993).
11. L. Castagnoli, C. Vetriani, G. Cesareni, *J. Mol. Biol.* **237**, 378 (1994).
12. P. F. Predki and L. Regan, *Biochemistry* **34**, 9834 (1995).
13. G. MacBeath, P. Kast, D. Hilvert, *Protein Sci.* **7**, 325 (1998).
14. G. MacBeath, P. Kast, D. Hilvert, unpublished results.
15. K. E. Neet and D. E. Timm, *Protein Sci.* **3**, 2167 (1994).
16. C. J. Bult *et al.*, *Science* **273**, 1058 (1996).
17. For comparison, k_{cat}/k_{uncat} for EcCM under these conditions is 3×10^6 [calculated from the activation parameters reported in C. C. Galopin, S. Zheng, D. B. Wilson, B. Ganem, *Tetrahedron Lett.* **37**, 8675 (1996); *ibid.* **38**, 1467 (1997)]. The efficiency of MjCM at 85°C, the optimal growth temperature for *M. jannaschii*, actually decreases somewhat ($k_{cat}/k_{uncat} = 9 \times 10^4$) because of the steeper temperature dependence of the uncatalyzed reaction.
18. V. Biou, A. Yaremchuk, M. Tukalo, S. Cusack, *Science* **263**, 1404 (1994).
19. M. G. Oakley and P. S. Kim, *Biochemistry* **36**, 2544 (1997).
20. Libraries were constructed in plasmid pKMCMT-W, which features the chorismate mutase gene (*aroQ$_f$*) from *M. jannaschii* under control of the *trc* promoter. Duplication of Leu^{20}-Lys^{21}, introduction of a Hind III site, and mutation of Leu^{20a} to Glu and Ile^{77} to Arg was accomplished by use of the polymerase chain reaction (PCR) with template pKMCMT-W and primers MNEG-S (5′-ATAAGATACTTAAGGAAAAGCTTATTGCTGAAAGAAATAGTTTAGCT-3′) and MNEG-N (5′-CTTTTGGAGGGCCTTATTATGCTCTATAAGACGTTGAAATATTTTAATGCCAAT-3′). Cloning of this product into pKMCMT-W with restriction enzymes Afl II and Eco 0109I gave pKMCMT-NEG. Six random codons were then introduced into the modified *aroQ$_f$* using the PCR with template pKMCMT-W and primers SEQT-S (5′-GTGAGCGGATAACAATTTCACA-3′) and MJMA-N (5′-TCAGCAATAAGCTTTTCSNNSNNSNNSNNSNNSNNCTTAAGGATCTTATTGTCAATCTCATCAATCT-3′) (N = A, C, G, or T; S = C or G). Cloning of this library of products into pKMCMT-NEG with restriction enzymes Nde I and Hind III gave library pKMCMT-LibA.
21. P. Kast, M. Asif-Ullah, N. Jiang, D. Hilvert, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5043 (1996).
22. The host strain for protein production was KA13, a CM-deficient *E. coli* with an isopropyl β-D-thiogalactopyranoside–inducible T7 RNA polymerase gene integrated into the chromosome.
23. R. Piccoli *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 1870 (1992).
24. Sedimentation equilibrium runs were performed from 3000 to 17,000 revolutions per minute with 2, 10, and 30 μM protein concentrations in phosphate-buffered saline (PBS) at 20°C. Partial specific volumes and solvent densities were calculated with the program Sednterp [T. M. Laue, B. D. Shah, T. M. Ridgeway, S. L. Pelletier, in *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, S. E. Harding, A. J. Rowe, J. C. Horton, Eds. (Royal Society of Chemistry, Cambridge, England, 1992), pp. 90–125] and data were analyzed by nonlinear least-squares fitting with the Origin software provided by Beckman.
25. M. J. Bennett, M. P. Schlunegger, D. Eisenberg, *Protein Sci.* **4**, 2455 (1995).
26. A. Di Donato, V. Cafaro, G. D'Alessio, *J. Biol. Chem.* **269**, 17394 (1994).
27. S. Trinkl, R. Glockshuber, R. Jaenicke, *Protein Sci.* **3**, 1392 (1994).
28. Supplementary material is available at www.sciencemag.org/feature/data/976438.shl.
29. Reaction kinetics were measured with 50 nM enzyme at 20°C in PBS, supplemented with bovine serum albumin (0.1 mg/ml). Inclusion of bovine serum albumin did not alter the aggregation state of the protein as verified by analytical size-exclusion column chromatography. Initial rates were determined by monitoring the disappearance of chorismate spectrophotometrically at 274 nm ($\varepsilon_{274} = 2630\ M^{-1}\ cm^{-1}$). Steady-state kinetic parameters k_{cat} and K_m were calculated from the initial rates.
30. M. J. Bennett, S. Choe, D. Eisenberg, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3127 (1994).
31. Removing six amino acids from the carboxyl terminus of MjCM increased its resistance to proteolytic cleavage without affecting the properties of the protein. Consequently, all detailed biochemical and biophysical characterizations of mMjCM were also carried out on a truncated derivative, which behaved identically to the full-length mMjCM in all aspects tested.

32. We thank H. Lashuel and J. Kelly for analytical ultracentrifugation studies. G.M. is the recipient of a Natural Sciences and Engineering Research Council of Canada 1967 Centennial Postgraduate Scholarship and an Eli Lilly Graduate Student Fellowship. Supported by the Skaggs Institute for Chemical Biology at The Scripps Research Institute.

1 December 1997; accepted 2 February 1998

Dependence of BSAP Repressor and Activator Functions on BSAP Concentration

Jeffrey J. Wallin,* Edwin R. Gackstetter, Marian Elliott Koshland†

During a B cell immune response, the transcription factor BSAP maintains its activator functions but is relieved of its repressor functions. This selective targeting of BSAP activities was shown to be regulated by a concentration-dependent mechanism whereby activator motifs for BSAP had a 20-fold higher binding affinity than repressor motifs. An exchange of activator and repressor motifs, however, showed that the context of the motif, rather than the affinity, determined whether BSAP operated as an activator or repressor.

BSAP, the transcription factor encoded by the *Pax5* gene, is a key regulator of B lymphocyte development (*1*). Its critical role in early B cell lymphopoiesis has been established by targeted disruption of the mouse *Pax5* gene (*2*). B cells from the mutant mice are blocked at the pro-B stage and do not develop further. BSAP also participates in the later antigen-driven stages of B cell differentiation. Analysis of the changes in gene expression induced by antigen and cytokine signals has shown that BSAP positively regulates the expression of the CD19 component of the B cell antigen coreceptor and germ-line transcription of the ε gene (Iε), which is a prerequisite for a switch to immunoglobulin E (IgE) synthesis (*3*, *4*). BSAP can also serve as a negative regulator during development. During an immune response, BSAP blocks are removed from the heavy chain 3′ enhancer to allow isotype switching and from the J chain gene promoter to allow synthesis and secretion of pentamer IgM antibody (*5*–*7*).

Because the positive functions of BSAP are maintained at the same time that its negative functions are relieved, we investigated how these different regulatory activities are accomplished. BSAP concentrations are high in pre-B, immature B, and mature B cells; once an antigen signal is delivered, however, the concentrations progressively decrease from the presecretor to the mature plasma cell stage, in which BSAP is almost undetectable (*7*, *8*). The decrease in BSAP expression is induced by signals from interleukin-2 (IL-2) or IL-5, which can down-regulate BSAP RNA expression to 20 to 25% of untreated cells (*7*, *9*). Overexpression of BSAP in plasma cell lines increases BSAP repressor activity: J-chain RNA and secreted IgA decrease (*7*, *10*). This expression pattern suggested that BSAP activities might be selectively targeted during antigen-driven differentiation by competition among the BSAP sites for the diminishing amounts of factor (*7*, *11*).

We obtained experimental evidence for a concentration-dependent regulation by determining the relative affinities of BSAP activator and repressor elements from gel-shift assays (Fig. 1). BSAP binding to a consensus recognition sequence (*12*) was competed by adding increasing amounts of the BSAP recognition motifs with estab-

Immunology Division, Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720, USA.

*To whom correspondence should be addressed.
†Deceased.