

All-codon scanning identifies p53 cancer rescue mutations

Roberta Baronio¹, Samuel A. Danziger^{1,2}, Linda V. Hall¹, Kirsty Salmon¹,
G. Wesley Hatfield^{1,3,4}, Richard H. Lathrop^{1,2,5,*} and Peter Kaiser^{6,*}

¹Institute for Genomics and Bioinformatics, ²Department of Biomedical Engineering, ³Department of Microbiology and Molecular Genetics, ⁴Department of Chemical Engineering and Materials Science, ⁵Department of Computer Science and ⁶Department of Biological Chemistry, University of California, Irvine, CA 92697, USA

Received April 16, 2010; Revised June 4, 2010; Accepted June 7, 2010

ABSTRACT

In vitro scanning mutagenesis strategies are valuable tools to identify critical residues in proteins and to generate proteins with modified properties. We describe the fast and simple All-Codon Scanning (ACS) strategy that creates a defined gene library wherein each individual codon within a specific target region is changed into all possible codons with only a single codon change per mutagenesis product. ACS is based on a multiplexed overlapping mutagenesis primer design that saturates only the targeted gene region with single codon changes. We have used ACS to produce single amino-acid changes in small and large regions of the human tumor suppressor protein p53 to identify single amino-acid substitutions that can restore activity to inactive p53 found in human cancers. Single-tube reactions were used to saturate defined 30-nt regions with all possible codon changes. The same technique was used in 20 parallel reactions to scan the 600-bp fragment encoding the entire p53 core domain. Identification of several novel p53 cancer rescue mutations demonstrated the utility of the ACS approach. ACS is a fast, simple and versatile method, which is useful for protein structure–function analyses and protein design or evolution problems.

INTRODUCTION

A simple and efficient approach that scans across a defined gene region and systematically produces all

possible single-codon mutations within that region would be highly desirable for many applications. Random mutagenesis, coupled with genetic screening or selection strategies to identify desired phenotypes, is often used to isolate proteins with modified activities or to identify critical residues in proteins (1–3). Protein-coding DNA regions are typically mutagenized *in vitro* using codon cassette mutagenesis, site-directed mutagenesis, chemical modification of nucleotides, or error-prone PCR conditions (4–7). Error-prone PCR-based strategies are biased towards certain nucleotides or nucleotide combinations, rarely produce two or three base changes in a single codon, typically do not achieve saturation, may introduce multiple amino acid changes per product, and cannot be controlled in detail. Codon cassette mutagenesis requires specific sequence conditions and targets a single codon at a time, thus requiring considerable effort if saturation of extended gene regions is desired (5,8). More directed approaches, such as site-directed mutagenesis or its extensions to alanine scanning (9) and iterative saturation mutagenesis (10), sample only selected residues and are labor-intensive for saturation mutagenesis of large gene regions or whole genes. The All-Codon Scanning (ACS) strategy presented here overcomes many of the above restrictions. ACS allows the systematic change of each codon in a defined gene region into all other codons, while yielding only a single changed codon per mutagenesis product.

ACS mutagenesis of small and large regions of the human tumor suppressor gene *p53* was used here to identify amino acid changes that can restore activity of p53 mutants found in human cancer. The *p53* gene encodes a tumor suppressor protein that is a key cellular defense against cancer (11). *p53* mutations occur in ~50% of human cancers, and about three-quarters of those

*To whom correspondence should be addressed. Tel: +1 949 824 9442; Fax: +1 949 824 2688; Email: pkaiser@uci.edu
Correspondence may also be addressed to Richard H. Lathrop. Tel: +1 949 824 4021; Fax: +1 949 824 4056; Email: rickl@uci.edu
Present address:

Kirsty Salmon, Verdezyne, Inc., Carlsbad, CA 92010, USA

mutations are single-point missense mutations in the *p53* core domain (12,13). Restoring *p53* activity in human tumors is therefore theoretically possible. Indeed, reactivated *p53* holds great therapeutic promise because animal models have shown that reintroduction of active *p53*, even in advanced tumors, leads to tumor regression (14–16). As demonstrated *in vivo*, *p53* cancer mutants can be reactivated through intragenic second-site suppressor mutations (17–19). These so-called ‘cancer rescue’ mutations suggest that drugs mimicking suppressor mutations might be designed (20). Therefore, studies on cancer rescue mutations aim to identify regions in *p53* where perturbations may cause *p53* reactivation. Cancer rescue mutations may also influence protein-protein interactions or protein stability without necessarily inducing major structural disruptions. Thus, identification of many or all *p53* cancer rescue mutations could provide a better structure–function understanding of the phenomenon of *p53* reactivation and support future drug development efforts.

As described below, the ACS strategy can be combined with synthetic DNA technologies to optimize hybridization fidelity. Several computational approaches exist that allow the design of synthetic DNA molecules and encompass a number of DNA sequence-determined properties without changing the encoded protein sequence (21–25). This is of course possible because of the degeneracy of the genetic code, and allows simultaneous optimization for several arbitrarily specified sequence properties, such as its own self-assembly, codon usage (26), GC/AT ratio (27,28), translational kinetics (29), codon-pair bias (30–32) and other DNA structural scales (33). Particularly useful for the ACS strategy is the Computational Optimized DNA Assembly (CODA) system (34), which exploits the degeneracy of the genetic code to perform multiple sequence optimizations, including removal of deleterious RNA cross-hybridizations and secondary structures. The resulting global thermodynamic optimization of the DNA sequence implies that every location in the gene is assigned a globally unique thermodynamic address. However, any modern gene design software that removes RNA cross-hybridizations and secondary structure may be used instead of CODA.

The ACS strategy described here is a quick and simple method that overcomes several shortcomings of current scanning mutagenesis approaches. ACS creates a sequence defined gene library with (i) each individual codon in a specific target region changed to all other possible codons, and (ii) only a single codon change per mutagenesis product.

MATERIALS AND METHODS

Materials

Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA, USA). PfuUltra II ultra DNA Polymerase was purchased from Stratagene (La Jolla, CA, USA). Restriction enzymes and T4 DNA ligase were obtained from New England Biolabs (Ipswich, MA, USA).

ACS mutagenesis

A pool of 10 forward and 10 reverse oligonucleotides, each with a single randomized NNN codon at its center was used in each mutagenesis reaction. NNN yields all possible codons and so reduces the possibility of codon-level effects that might interfere with translation or expression (e.g. rare codons or bad codon pairs), while the stop codons are inactive in our genetic selection and so were not seen. The ACS method can be modified easily to use NNK or NNS (or any desired triplets) for use where there was less concern about codon-level effects or stop codons would be deleterious. Similarly, in applications where DNA synthesis bias would be a concern, custom preparation of oligonucleotides with balanced base ratios is available (<http://www.idtdna.com/catalog/analytical/Page1.aspx>).

The constant flanking regions surrounding each NNN codon matched the target incorporation site and were extended at least four bases past any other NNN codon. Consequently, incorporation of any mutagenesis oligonucleotide into a product must mask any previously incorporated NNN, and so only one NNN codon is possible in any product. Mutagenesis oligonucleotide extension can potentially lead to a bias towards products that are generated from oligonucleotides with higher annealing temperatures. To overcome this problem we used the Stratagene online mutagenesis oligonucleotide designer (<http://www.stratagene.com/sdmdesigner/default.aspx>) to obtain roughly equal annealing temperatures greater than 68°C.

The 50 µl PCR reaction was carried out with 50 ng templates, 6 µM total primer concentration (300 nM of each forward and reverse primer), 400 µM dNTPs and 5 U of Pfu Ultra II polymerase. The reaction was initiated at 94°C for 10 min, followed by 18 cycles (30 s at 94°C, 40 s at 55°C and 200 s at 72°C), and concluded by incubation at 72°C for 15 min. The amplification products were digested with DpnI to remove the methylated template plasmids. At this point, a plasmid library containing completely saturated target regions of a gene of interest can be generated by transformation into *Escherichia coli*. Alternatively, the mutagenized products can be amplified by PCR and introduced into target vectors by homologous recombination or used in many other standard applications.

Yeast assays to identify and evaluate intragenic suppressor mutations

A region containing the *p53* core domain (amino acids 50–334) was re-amplified from the mutagenesis mix. The PCR products were purified (QIAquick™ PCR purification kit, Qiagen, Germany) and used to generate *p53*-expression plasmids in yeast using the gap-repair strategy (35). To identify mutations that reactivate *p53* cancer mutants we used a *p53*-tester yeast strain. Growth of this yeast strain in medium lacking uracil depends on active *p53* because the *URA3* gene is engineered to be under transcriptional control of *p53* (18,36). The plasmid pTW300, which is a centromeric plasmid expressing human *p53* under control of the *ADHI*

promoter (36), was gapped with AgeI/StuI. These restriction sites are positioned so that the ACS final PCR products overlapped with both ends of the gapped plasmids by at least 60 nt at both sites. Taking advantage of the highly efficient homologous recombination in *Saccharomyces cerevisiae*, 500 ng of the PCR products and 500 ng of the gapped plasmid were co-transformed into yeast strain RB379. The p53-tester yeast strain RB379 was engineered so that expression of the *URA3* gene is dependent on active p53 (18). Cells were directly plated onto SC plates lacking uracil and incubated at 37°C, thus selecting for repaired, functional plasmids and a possibly functional p53 molecule. These plasmids were directly sequenced from yeast cells, after colony-PCR amplification (Genewiz, San Diego, CA, USA).

CODA optimization and assembly of the full length p53 gene

The p53 gene sequence was CODA-optimized to encode its own correct self-assembly as described (34). Computations were performed on a 64-node cluster of 3 GHz Xeon dual processors. Computer code is available upon request. The p53 sequence was also designed for optimal expression in *S. cerevisiae* using synonymous codon substitutions. Codon usage for yeast was estimated using the codon adaptation index (37) and codon pair statistics (30–32,38). The CODA output is a list of DNA oligonucleotide sequences and a set of instructions for combining the purchased oligonucleotides by polymerase extension and PCR into the full length synthetic gene.

For each overlapping intermediate DNA fragment, the constituent oligonucleotide set was mixed to final concentrations of 0.1 μM with an excess (1 μM) of the 5' and 3' primer oligonucleotide. Each oligonucleotide set was extended to an intermediate DNA fragment by a PCR reaction using 2.5 U of PfuUltra II ultra DNA polymerase (Stratagene), 200 μM dNTPs (Roche Diagnostics, Indianapolis, IN) and 1X PfuUltra reaction buffer. These reactions were performed in a MJ Research PTC-225 thermal cycler using the following calculated-control protocol: 10 min denaturation step at 95°C, followed by 35 cycles of 30 s at 95°C, 30 s at 68°C and 20 s at 72°C, and a final step of 5 min at 72°C. The PCR products were subjected to electrophoresis in a 1% agarose gel, PCR purified (Quiaquick PCR clean up, Quiagen, Valencia, CA, USA), and quantified by NanoDrop (Thermo Scientific, Wilmington, DE, USA). Each overlapping intermediate DNA fragment was mixed to a final concentration of 0.1 μM along with 1 μM NdeI/BamHI cloning oligonucleotides, 2.5 U of PfuUltra II ultra DNA polymerase (Stratagene), 200 μM dNTPs (Roche Diagnostics, Indianapolis, IN), and 1× PfuUltra reaction buffer. These primer extension and PCR amplification reactions were performed in a MJ Research PTC-225 thermal cycler as above, but with 45 s extension at 72°C. The full-length p53 gene was separated in a 1% agarose gel, PCR purified, and cloned into a pCODA vector. The sequence was confirmed by DNA sequencing.

Cell culture and p53 evaluation methods

H1299 and Saos-2 cell lines (p53 negative) were grown in high-glucose DMEM with 10% fetal bovine serum. All transfections were done with Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. All experiments were done in triplicate. p53-dependent reporter activity values were normalized to *Renilla* luciferase activity by co-transfection with expression plasmids carrying the *Firefly* luciferase gene under the control of the *p21* promoter and a *CMV* promoter driven p53 (36). Cells were harvested 48 h after transfection and analyzed for luciferase activity with the Dual-Glo Luciferase Assay System (Promega, Madison, WI, USA) according to the manufacturer's protocol. Light emission was detected with the Monolight 2010 luminometer (Analytical Luminescence Laboratory, San Diego, CA, USA).

Three dimensional visualization of p53 mutants

The location of cancer mutants and cancer rescue mutants in the p53 3D structure was visualized using Visual Molecular Dynamics (VMD) (<http://www.ks.uiuc.edu/Research/vmd/>).

RESULTS

Self-assembly of CODA-optimized human p53

A CODA-optimized human p53 core domain DNA sequence was optimized to code for the same amino acid sequence as the native gene, to be easily self-assembled from synthetic oligonucleotides by polymerase extension, respect hypothesized codon usage and codon pair preferences (30–32) in yeast, and especially to avoid RNA secondary structure and off-target cross-hybridization, which is known to be a source of mutagenesis failure (33). Nucleotide changes required to achieve these properties (Figure 1A) were computed by the CODA algorithm as described (34).

Briefly, the p53 gene was divided into six overlapping intermediate DNA fragments, each of which was subdivided into 10 short overlapping and abutting oligonucleotides. There exists a temperature gap within which, with high thermodynamic probability, correct hybridizations are annealed and incorrect hybridizations are melted. The melting temperature distributions of the CODA-designed p53 sequence show a melting temperature gap of 20°C (Figure 1B), which essentially results in a globally unique thermodynamic address for each location in the target gene and thereby minimizes mutagenesis difficulties due to miss-priming and cross-hybridization. Self-assembled overlapping intermediate DNA fragments were obtained by primer extension and PCR amplification. The overlapping intermediate DNA fragments were combined in another PCR amplification reaction to generate a single product, sequence verified, full-length p53 core domain (Figure 1C).

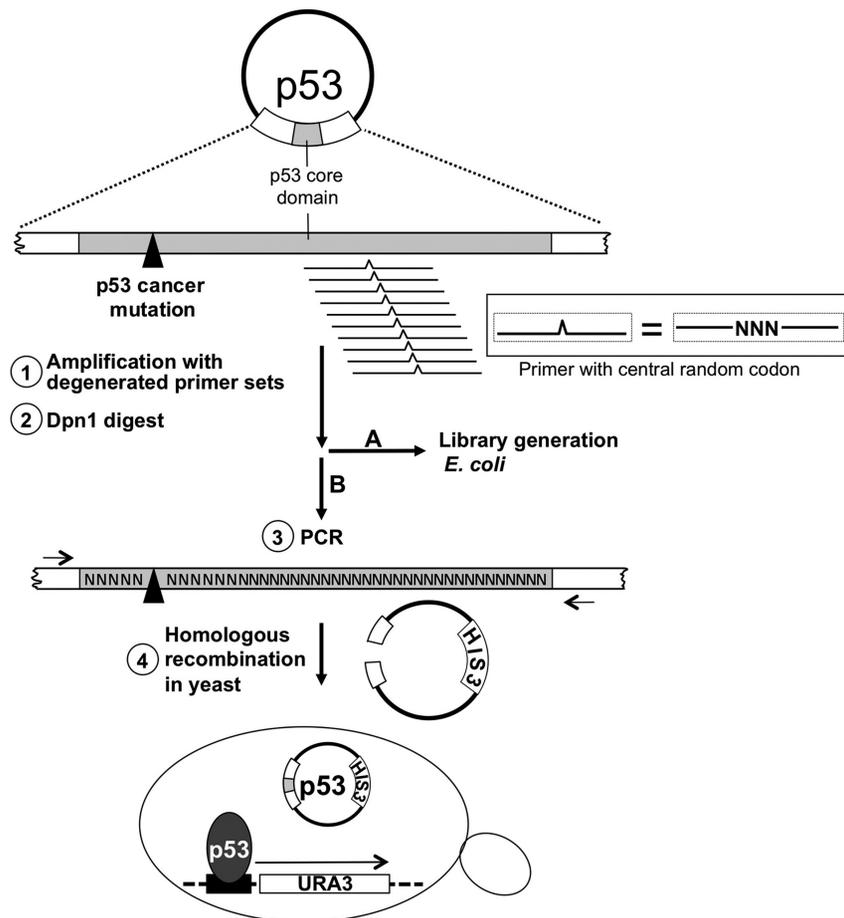


Figure 2. Schematic representation of the ACS strategy and the p53 cancer mutant rescue screen. A set of overlapping, degenerate oligonucleotides with roughly equal initial annealing temperatures was designed to change each single codon in a 30 bp wide region. Parallel reactions can be used to mutagenize large regions or entire genes. Parent plasmids were removed by Dpn1 digestion. At this point, a plasmid library containing completely saturated target regions of a gene of interest could be generated by transformation into *E. coli*. Alternatively, the mutagenized products could be amplified by PCR and introduced into target vectors by homologous recombination, or be used in many other standard applications. To select intragenic mutants that restore activity to p53 mutants found in human cancer, the p53 core domain was amplified from the mutagenized pool and introduced into a p53-tester yeast strain by homologous recombination. Yeast cells expressing re-activated p53 were identified by uracil prototrophy.

First, we verified by sequencing more than 500 randomly chosen individual mutagenesis products that, at a given amino acid position, all 64 possible codon insertions were observed, and all 20 amino acids were represented at approximately their expected frequencies (correlation coefficient $r = 0.79$) (Figure 3). We observed a very high mutagenesis efficiency (>95%), which we attribute to the unique thermodynamic profile of the synthetic gene design. The melting temperature distributions of the CODA-designed p53 sequence show a melting temperature gap of 20°C (Figure 1B). Thus, correct annealing of any given oligonucleotide was strongly favored over hybridization at an incorrect position. Even with a 3-nt mismatch mutagenesis oligonucleotide, the temperature gap between correct hybridizations and incorrect hybridizations remained at a very selective 15°C. This high thermal gap avoids annealing to incorrect gene locations and solves one of the main obstacles to efficiently produce sequence-defined multiplexed DNA libraries (33).

Next, we chose the region encoding amino acids 232–241 of p53 to test the efficiency of ACS mutagenesis in identification of rescue mutations for the p53^{R158L} cancer mutant. Amino acid region 232–241 in mutant p53^{R158L} was specifically selected because it had been extensively explored previously by error-prone PCR and was identified as a ‘global suppressor motif’ due to the accumulation of cancer rescue mutations (36). The region containing the p53 core domain was re-amplified from the mutagenesis mix. The PCR products were purified and used to generate p53-expression plasmids in yeast using the gap-repair strategy (Figure 2). As described in detail in the ‘Materials and Methods’ section, a p53-tester yeast strain with p53-controlled expression of the selectable *URA3* gene was used for these experiments to identify mutations that restore p53 activity (36).

For each ACS reaction we obtained at least 200 000 transformants, ensuring full coverage of all generated mutants. Specifically, considering mutations at 10 possible codon locations, each of which can have any

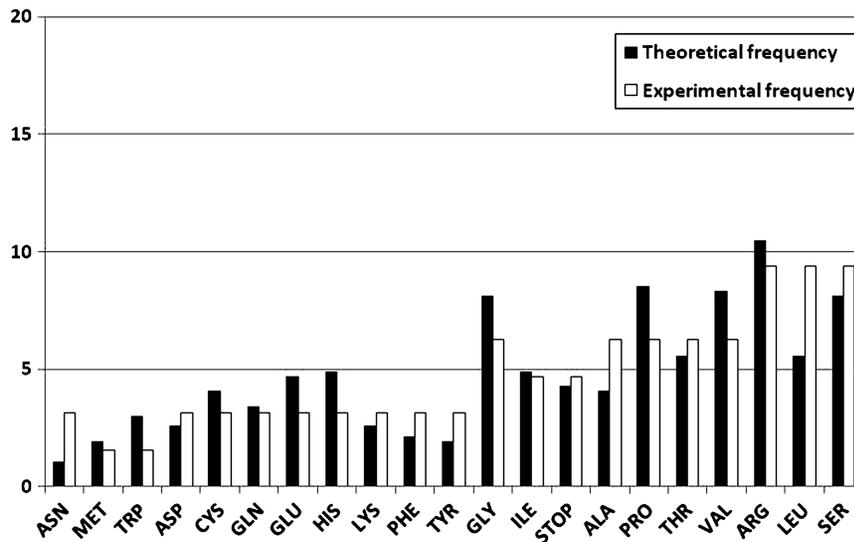


Figure 3. Representation of the experimental and theoretical frequency of codons in position 239 obtained by ACS.

one of 64 codons, the 640 theoretical different mutants are covered by >200 000 transformants with >300-fold coverage (as >200 000/640). Such high coverage almost certainly guarantees that every codon change was represented ($P < 1.0e - 130$ by the binomial distribution with Bonferoni correction for 640 trials).

ACS mutagenesis of the *p53* region encoding amino acids 232–241 identified all five previously described rescue mutants and one new mutation (Y234L) that reactivated the $p53^{R158L}$ cancer mutant (Table 1). Interestingly, the Y234L mutation required a minimum of 2 nt changes and so is unlikely to be found using error-prone PCR, thus providing evidence for the value of our exhaustive approach.

ACS identified *p53* cancer rescue mutations in amino acid region 114 to 123

We next wanted to scan a novel region in *p53* for reactivating second-site mutations that can restore activity to *p53* mutants. We considered the 10 *p53* cancer mutants that are most commonly found in human cancer and can be constructed so that they differ by two or more nucleic acid changes from the wild-type (R249S, G245S, H179R, R273H, R248L, R158L, R280T, P151S, P152L, P278L). Cancer mutants were restricted to those with multiple nucleotide changes from wild-type *p53* to minimize false-positive errors due to spontaneous reversion, although analysis of the experimental results later indicated that this precaution was not necessary. These 10 cancer mutants were used for ACS mutagenesis of amino acid region 114–123. Note that one of these *p53* mutants ($p53^{R158L}$) was used in the ACS experiment targeting region 232–241. Region 114–123 was considered a potential cancer rescue region because multiple amino acid changes occurred there spontaneously in previous cancer rescue mutant screens but no single amino acid change cancer rescue mutations had been found previously [(36); Brachmann, R. K., personal communication]. In addition, a structure-based computational classifier trained to

Table 1. Summary of *p53* cancer rescue mutants obtained by ACS targeting amino acid region 232–241 of *p53* cancer mutant $p53^{R158L}$

<i>p53</i> Cancer mutant	Rescue mutation
R158L	N239Y
R158L	N239F
R158L	N239R
R158L	N235K
R158L	Y234V
R158L	Y234L*

Asterisks label rescue mutants that required at least 2-nt changes for that amino-acid substitution. New rescue mutants are in boldface.

predict regions in *p53* that likely contain cancer rescue mutations selected region 114–123 (39). Therefore, this region was thought likely to include single amino acid cancer rescue mutants and was selected for ACS.

We obtained a minimum of 200 000 yeast transformants for each library, which represents >300-fold coverage of the expected 640 mutants in each library. We isolated a total of 146 transformants that grew on plates lacking uracil. The *p53* core domains of all 146 potential rescue mutants were amplified by colony PCR and sequenced. Three of the isolated uracil prototroph strains contained wild-type *p53*, which was likely a result of reversion of the cancer mutation due to selection pressure. However, the vast majority of sequences (143) retained the cancer mutation. A total of 13 different potential rescue mutants showed both the cancer mutation and a second mutation that changed the amino acid sequence in the mutagenized region. All potential rescue mutants were regenerated using site directed mutagenesis, transformed into the *p53*-tester yeast strain, and analyzed for their ability to confer uracil prototrophy (Figure 4). All rescue mutations were confirmed. Six of them had a strong reactivation activity, whereas seven showed only a very small gain in activity compared to the cancer mutant and were not further tested (Table 2). We measured *p53* activity of the strong rescue mutants in

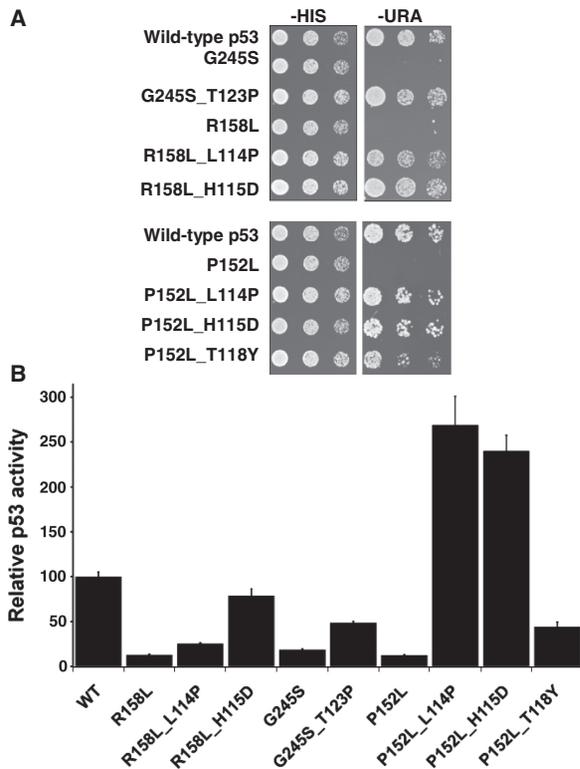


Figure 4. Reactivation of p53 cancer mutants by second site mutations in amino acid region 114–123. **(A)** Yeast cells expressing the *URA3* gene from a p53-dependent promoter were transformed with centromeric plasmids (*HIS3* selection marker) expressing wild-type human p53, or the mutants indicated, under control of the *ADHI* promoter. Serial dilutions of cells grown to mid-log phase in medium lacking histidine were spotted onto plates either lacking histidine or uracil. Growth on plates lacking uracil is dependent on expression of the *URA3* gene and is a measure of p53 activity. Mutants are named with the cancer mutation appearing first followed by the putative cancer rescue mutation. **(B)** p53 transcriptional activity was evaluated in transient reporter gene assays in H1299 cells, which lack endogenous p53 activity. Firefly luciferase activity, adjusted for transfection efficiency with *Renilla* luciferase activity, was determined 48 h after transfection. The activity of wild-type p53 was set to 100%. SD were calculated from at least three independent experiments.

Table 2. Summary of p53 cancer rescue mutants obtained by ACS targeting amino acid region 114–123 of 10 p53 cancer mutants commonly found in human tumors (R249S, G245S, H179R, R273H, R248L, R158L, R280T, P151S, P152L, P278L)

p53 Cancer mutant	Rescue mutation
G245S	T123P
R158L	L114P*
R158L	H115D
R158L	L114P*
P152L	H115D
P152L	T118Y*
p53 Cancer mutant	Weak Rescue mutation
R158L	T123S
R158L	H115N*
P152L	T123G*
P152L	T123A
P152L	T123P*
P152L	T123S
P152L	H115E*

Asterisks and boldface are as in Table 1.

human H1299 cells, which lack endogenous p53 activity. p53 cancer mutants and rescue mutants were expressed in H1299 cells and activity was measured using a p53-dependent luciferase reporter (18) (Figure 4). Without exception, mutations confirmed in the yeast-based screen also reactivated p53 cancer mutants in human cells.

All p53 cancer rescue mutants identified in amino acid region 114–123 were new and had not been identified by previous error-prone PCR mutagenesis approaches using the same yeast-based p53 activity assay. ACS resulted in a mutant that reactivated the p53 cancer mutant p53^{P152L}, which had not been rescued before. Importantly, three out of the six identified amino acid changes that strongly reactivated p53 cancer mutants required 2- or 3-nt changes. Such mutations are underrepresented in current *in vivo* or *in vitro* mutagenesis approaches, but were easily identified using our complete and unbiased exhaustive approach.

Combined ACS reactions to scan larger domains or entire genes

ACS reactions easily can be performed in parallel to produce gene libraries that target much larger gene regions or entire genes. We used ACS to saturate all 200 amino acids of the p53 core domain with all possible single amino acid changes. The ACS method consists of single-tube reactions, each covering a region of 30 bases. Twenty parallel ACS reactions were performed and the mutagenesis products were introduced to the p53-tester yeast strain (Figure 2). Combining these single ACS reactions produced a completely saturated library with exactly one mutation in each product. At the same time, the modularity of the system allows to avoid the modification of specific gene areas of interest or even key residues, such as the p53 cancer mutation itself.

The three p53 cancer mutants p53^{R175H}, p53^{R273H} and p53^{G245S} were exhaustively scanned for single-change cancer rescue mutants throughout the 200 amino acid core region. We selected these particular cancer mutants due to their specific characteristics. p53^{R175H} is the most common cancer mutant found in human tumors, but reversion to wild-type requires only a single nucleotide change and we were concerned about a potentially high spontaneous reversion rate interfering with ACS screening. Eventually, however, random reversion to wild-type p53 turned out to be insignificant. We also chose to analyze p53^{R273H}, the most common cancer mutant that differs by 2-nt changes from wild-type p53, to avoid altogether the possibility of spontaneous reversion. Neither cancer mutant had been rescued previously by intragenic single amino acid changes, and thus might have been too stringent an initial test case for ACS. Therefore, we also analyzed p53^{G245S}, the most common cancer mutant that had been rescued previously by a single amino acid change.

Table 3 summarizes the results of ACS saturation mutagenesis on the 200 amino acids in the p53 core domain for these three different cancer mutants. p53^{R175H} could not be reactivated by a single amino acid change. For p53^{R273H} and p53^{G245S}, we recovered all previously

Table 3. Summary of p53 cancer rescue mutants obtained by ACS targeting the 600-nt region encoding the p53 core domain of cancer mutants p53^{R273H} and p53^{G245S}

p53 Cancer mutant	Rescue mutation
G245S	N239F
G245S	F113L
G245S	S240Y
G245S	T123P
G245S	N239Y
R273H	S240R*
R273H	S240Q*
R273H	C242R
R273H	N263V*

Asterisks and boldface are as in Table 1.

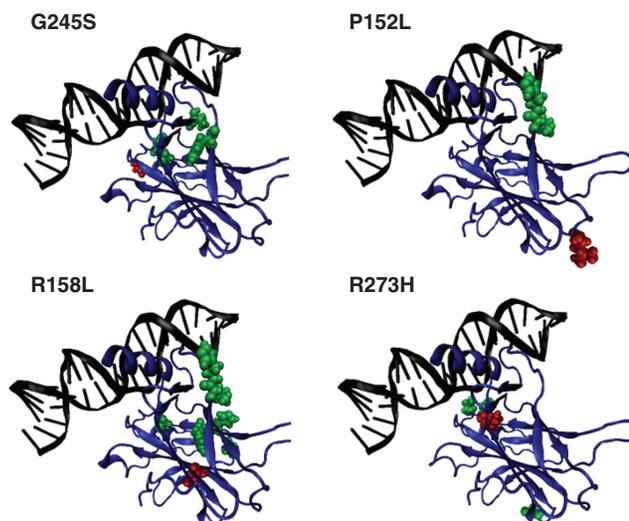
known single rescue mutations and identified six new re-activation sites (Table 3). Three of the six new sites required two or three changes from wild-type p53, and so are unlikely to be seen using error-prone PCR. These results further demonstrate the utility of ACS for efficient, economic, and systematic saturation scanning mutagenesis in defined gene sections.

DISCUSSION

In summary, the ACS strategy creates a sequence defined gene library, which changes each individual codon in a specific target region into all possible codons, with only a single codon change per mutagenesis product. The ACS approach is fast, simple, and generally applicable to any problem that requires systematic changes of defined regions in a DNA sequence, including protein design, directed evolution, structure/function analyses and promoter studies.

Individual ACS mutagenesis reactions can target changes in a single codon (or even a single base pair position) for as many as 30 contiguous base pairs. The upper base pair target limit restricts primer length to avoid incorporation of mutations due to primer synthesis errors, considering a typical error rate of one per 600 synthesized nucleotides (40–43). By performing parallel reactions, mutagenesis can be expanded to larger regions or whole genes. The ACS system allows complete control over the target regions; specific gene sections, or even specific nucleotides, easily can be protected from mutagenesis. Thus, ACS is an adaptable and versatile tool.

Here, ACS was used to mutagenize small and large regions of the human tumor suppressor gene *p53*. Eleven common p53 cancer mutants, which collectively are found in well over one million diagnosed cancer cases per year, were tested. Reactivating intragenic mutations for four of them were identified. Moreover, the p53 cancer mutants p53^{P152L} and p53^{R273H} had not been reactivated previously by single amino acid changes. Thus, ACS obtained more complete biological information than previously had been obtained using methods such as error-prone PCR (Supplementary Table S1). Notably, 7 out of the

**Figure 5.** Three-dimensional location of cancer mutants and their rescue mutations. The 3D structure of p53 (51) was visualized with Visual Molecular Dynamics (52). Positions of cancer mutations are indicated in red and positions of rescue mutations are in green.

12 strong new p53 cancer rescue mutants we identified in this study required 2-or 3-nt changes, confirming the utility of ACS as an unbiased approach.

To understand better the regions selected and their relationship to the p53 protein, it is helpful to consider molecular visualizations of p53 (Figure 5). Cancer mutations are indicated in red and rescue mutations in green. The 3D locations are consistent with the hypothesis that different p53 cancer rescue mutants require different rescue mechanisms (44–46), and suggest that different rescue regions may correspond to distinct modes of cancer mutant reactivation. For example, rescue mutations for the p53 cancer mutant p53^{R273H} are located in distinct regions that do not overlap with rescue regions associated with other cancer mutants (Figure 5). A comprehensive identification of such rescue regions using ACS or similar approaches may contribute to targeted drug development, as it helps to identify surface exposed, drug accessible sites on p53 where perturbations may result in reactivation.

A significant proportion of human tumors express full length but mutated, inactive p53. Consequently, there are ongoing efforts to find small molecule drugs that mimic the cancer rescue effect of reactivating p53 second-site suppressor mutations (44,45,47,48). Despite impressive progress, efforts are hampered by a limited understanding of the p53 mutation–structure/function relationship (13,46,49,50). The ACS strategy described here has expanded and diversified the collection of rescue mutations that reactivate p53 cancer mutants. We hope that these efforts will lead to further insight into general structural changes that can rescue p53 cancer mutants, and thereby facilitate rational drug design approaches that seek to exploit similar effects.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Lydia Ho, Faezeh Salehi and Chris Wassman for helpful discussion and technical assistance. S. A. D. acknowledges support from the National Institutes of Health Biomedical Informatics Training Program. Materials and computer code will be provided under the terms of the standard Materials Transfer Agreement of the Regents of the University of California.

FUNDING

National Institutes of Health BISTI (grant number CA-112560, partial); National Science Foundation (grant number IIS-0326037). National Institutes of Health Biomedical Informatics Training Program (grant number LM-07443 to S.A.D.). Funding for open access charge: National Institutes of Health BISTI (grant number CA-112560); National Science Foundation (grant number IIS-0326037).

REFERENCES

- Jackel,C., Kast,P. and Hilvert,D. (2008) Protein design by directed evolution. *Annu. Rev. Biophys.*, **37**, 153–173.
- Liu,W., Sheppeck,J.E. 2nd, Colby,D.A., Huang,H.B., Nairn,A.C. and Chamberlin,A.R. (2003) The selective inhibition of phosphatases by natural toxins: the anhydride domain of tautomycin is not a primary factor in controlling PP1/PP2A selectivity. *Bioorg. Med. Chem. Lett.*, **13**, 1597–1600.
- Yuen,C.M. and Liu,D.R. (2007) Dissecting protein structure and function using directed evolution. *Nat. Methods*, **4**, 995–997.
- Hogrefe,H.H., Cline,J., Youngblood,G.L. and Allen,R.M. (2002) Creating randomized amino acid libraries with the quikchange multi site-directed mutagenesis kit. *Biotechniques*, **33**, 1158–1160, 1162, 1164–1155.
- Kegler-Ebo,D.M., Polack,G.W. and DiMaio,D. (1996) Use of codon cassette mutagenesis for saturation mutagenesis. *Methods Mol. Biol.*, **57**, 297–310.
- Lai,Y.P., Huang,J., Wang,L.F., Li,J. and Wu,Z.R. (2004) A new approach to random mutagenesis in vitro. *Biotechnol. Bioeng.*, **86**, 622–627.
- Vartanian,J.P., Henry,M. and Wain-Hobson,S. (1996) Hypermutagenic PCR involving all four transitions and a sizeable proportion of transversions. *Nucleic Acids Res.*, **24**, 2627–2631.
- Kegler-Ebo,D.M., Docktor,C.M. and DiMaio,D. (1994) Codon cassette mutagenesis: a general method to insert or replace individual codons by using universal mutagenic cassettes. *Nucleic Acids Res.*, **22**, 1593–1599.
- Cunningham,B.C. and Wells,J.A. (1989) High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science*, **244**, 1081–1085.
- Reetz,M.T. and Carballeira,J.D. (2007) Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat. Protoc.*, **2**, 891–903.
- Vogelstein,B., Lane,D. and Levine,A.J. (2000) Surfing the p53 network. *Nature*, **408**, 307–310.
- Hollstein,M., Rice,K., Greenblatt,M.S., Soussi,T., Fuchs,R., Sorlie,T., Hovig,E., Smith-Sorensen,B., Montesano,R. and Harris,C.C. (1994) Database of p53 gene somatic mutations in human tumors and cell lines. *Nucleic Acids Res.*, **22**, 3551–3555.
- Kato,S., Han,S.Y., Liu,W., Otsuka,K., Shibata,H., Kanamaru,R. and Ishioka,C. (2003) Understanding the function-structure and function-mutation relationships of p53 tumor suppressor protein by high-resolution missense mutation analysis. *Proc. Natl Acad. Sci. USA*, **100**, 8424–8429.
- Sharpless,N.E. and DePinho,R.A. (2007) Cancer biology: gone but not forgotten. *Nature*, **445**, 606–607.
- Ventura,A., Kirsch,D.G., McLaughlin,M.E., Tuveson,D.A., Grimm,J., Lintault,L., Newman,J., Reczek,E.E., Weissleder,R. and Jacks,T. (2007) Restoration of p53 function leads to tumour regression in vivo. *Nature*, **445**, 661–665.
- Xue,W., Zender,L., Miething,C., Dickins,R.A., Hernandez,E., Krizhanovsky,V., Cordon-Cardo,C. and Lowe,S.W. (2007) Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. *Nature*, **445**, 656–660.
- Brachmann,R.K., Vidal,M. and Boeke,J.D. (1996) Dominant-negative p53 mutations selected in yeast hit cancer hot spots. *Proc. Natl Acad. Sci. USA*, **93**, 4091–4095.
- Brachmann,R.K., Yu,K., Eby,Y., Pavletich,N.P. and Boeke,J.D. (1998) Genetic selection of intragenic suppressor mutations that reverse the effect of common p53 cancer mutations. *EMBO J.*, **17**, 1847–1859.
- Otsuka,K., Kato,S., Kakudo,Y., Mashiko,S., Shibata,H. and Ishioka,C. (2007) The screening of the second-site suppressor mutations of the common p53 mutants. *Int. J. Cancer*, **121**, 559–566.
- Boeckler,F.M., Joerger,A.C., Jaggi,G., Rutherford,T.J., Vepintsev,D.B. and Fersht,A.R. (2008) Targeted rescue of a destabilized mutant of p53 by an in silico screened drug. *Proc. Natl Acad. Sci. USA*, **105**, 10360–10365.
- Gao,X., Yo,P., Keith,A., Ragan,T.J. and Harris,T.K. (2003) Thermodynamically balanced inside-out (TBIO) PCR-based gene synthesis: a novel method of primer design for high-fidelity assembly of longer gene sequences. *Nucleic Acids Res.*, **31**, e143.
- Stemmer,W.P. (1994) DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc. Natl Acad. Sci. USA*, **91**, 10747–10751.
- Tian,J., Gong,H., Sheng,N., Zhou,X., Gulari,E., Gao,X. and Church,G. (2004) Accurate multiplex gene synthesis from programmable DNA microchips. *Nature*, **432**, 1050–1054.
- Welch,M., Govindarajan,S., Ness,J.E., Villalobos,A., Gurney,A., Minshull,J. and Gustafsson,C. (2009) Design parameters to control synthetic gene expression in *Escherichia coli*. *PLoS ONE*, **4**, e7002.
- Welch,M., Villalobos,A., Gustafsson,C. and Minshull,J. (2009) You're one in a googol: optimizing genes for protein expression. *J. Roy. Soc. Interface*, **6(Suppl. 4)**, S467–S476.
- Grosjean,H. and Fiers,W. (1982) Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. *Gene*, **18**, 199–209.
- Breslauer,K.J., Frank,R., Blocker,H. and Marky,L.A. (1986) Predicting DNA duplex stability from the base sequence. *Proc. Natl Acad. Sci. USA*, **83**, 3746–3750.
- Lio,P. (2002) Investigating the relationship between genome structure, composition, and ecology in prokaryotes. *Mol. Biol. Evol.*, **19**, 789–800.
- Sorensen,C.M. and Pierce,C.W. (1986) Identification and characterization of a suppressor T cell hybridoma specifically inducible by L-glutamic acid60-L-alanine30-L-tyrosine10 (GAT). *J. Immunol.*, **137**, 1455–1461.
- Gutman,G.A. and Hatfield,G.W. (1989) Nonrandom utilization of codon pairs in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **86**, 3699–3703.
- Hatfield,D.L., Dudock,B.S. and Eden,F.C. (1983) Characterization and nucleotide sequence of a chicken gene encoding an opal suppressor tRNA and its flanking DNA segments. *Proc. Natl Acad. Sci. USA*, **80**, 4940–4944.
- Irwin,B., Heck,J.D. and Hatfield,G.W. (1995) Codon pair utilization biases influence translational elongation step times. *J. Biol. Chem.*, **270**, 22801–22806.
- Wassman,C.D., Tam,P.Y., Lathrop,R.H. and Weiss,G.A. (2004) Predicting oligonucleotide-directed mutagenesis failures in protein engineering. *Nucleic Acids Res.*, **32**, 6407–6413.
- Larsen,L.S., Wassman,C.D., Hatfield,G.W. and Lathrop,R.H. (2008) Computationally Optimised DNA Assembly of synthetic genes. *Int. J. Bioinform. Res. Appl.*, **4**, 324–336.
- Muhlrad,D., Hunter,R. and Parker,R. (1992) A rapid method for localized mutagenesis of yeast genes. *Yeast*, **8**, 79–82.
- Baroni,T.E., Wang,T., Qian,H., Dearth,L.R., Truong,L.N., Zeng,J., Denes,A.E., Chen,S.W. and Brachmann,R.K. (2004)

- A global suppressor motif for p53 cancer mutants. *Proc. Natl Acad. Sci. USA*, **101**, 4930–4935.
37. Sharp, P.M. and Li, W.H. (1987) The codon Adaptation Index—a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.*, **15**, 1281–1295.
 38. Moore, G.L. and Maranas, C.D. (2002) Predicting out-of-sequence reassembly in DNA shuffling. *J. Theor. Biol.*, **219**, 9–17.
 39. Danziger, S.A., Baronio, R., Ho, L., Hall, L., Salmon, K., Hatfield, G.W., Kaiser, P. and Lathrop, R.H. (2009) Predicting positive p53 cancer rescue regions using Most Informative Positive (MIP) active learning. *PLoS Comput. Biol.*, **5**, e1000498.
 40. Baedeker, M. and Schulz, G.E. (1999) Overexpression of a designed 2.2kb gene of eukaryotic phenylalanine ammonia-lyase in *Escherichia coli*. *FEBS Lett.*, **457**, 57–60.
 41. Chalmers, F.M. and Curnow, K.M. (2001) Scaling up the ligase chain reaction-based approach to gene synthesis. *Biotechniques*, **30**, 249–252.
 42. Hoover, D.M. and Lubkowski, J. (2002) DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic Acids Res.*, **30**, e43.
 43. Withers-Martinez, C., Carpenter, E.P., Hackett, F., Ely, B., Sajid, M., Grainger, M. and Blackman, M.J. (1999) PCR-based gene synthesis as an efficient approach for expression of the A+T-rich malaria genome. *Protein Eng.*, **12**, 1113–1120.
 44. Bullock, A.N. and Fersht, A.R. (2001) Rescuing the function of mutant p53. *Nat. Rev. Cancer*, **1**, 68–76.
 45. Lambert, J.M., Gorzov, P., Veprintsev, D.B., Soderqvist, M., Segerback, D., Bergman, J., Fersht, A.R., Hainaut, P., Wiman, K.G. and Bykov, V.J. (2009) PRIMA-1 reactivates mutant p53 by covalent binding to the core domain. *Cancer Cell*, **15**, 376–388.
 46. Martin, A.C., Facchiano, A.M., Cuff, A.L., Hernandez-Boussard, T., Olivier, M., Hainaut, P. and Thornton, J.M. (2002) Integrating mutation data and structural analysis of the TP53 tumor-suppressor protein. *Hum Mutat*, **19**, 149–164.
 47. Bykov, V.J., Issaeva, N., Zache, N., Shilov, A., Hulterantz, M., Bergman, J., Selivanova, G. and Wiman, K.G. (2005) Reactivation of mutant p53 and induction of apoptosis in human tumor cells by maleimide analogs. *J. Biol. Chem.*, **280**, 30384–30391.
 48. Lain, S., Hollick, J.J., Campbell, J., Staples, O.D., Higgins, M., Aoubala, M., McCarthy, A., Appleyard, V., Murray, K.E., Baker, L. et al. (2008) Discovery, in vivo activity, and mechanism of action of a small-molecule p53 activator. *Cancer Cell*, **13**, 454–463.
 49. Cuff, A.L. and Martin, A.C. (2004) Analysis of void volumes in proteins and application to stability of the p53 tumour suppressor protein. *J. Mol. Biol.*, **344**, 1199–1209.
 50. Joerger, A.C., Ang, H.C., Veprintsev, D.B., Blair, C.M. and Fersht, A.R. (2005) Structures of p53 cancer mutants and mechanism of rescue by second-site suppressor mutations. *J. Biol. Chem.*, **280**, 16030–16037.
 51. Cho, Y., Gorina, S., Jeffrey, P.D. and Pavletich, N.P. (1994) Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science*, **265**, 346–355.
 52. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.