ARTICLE

Optimization of Protocol for Sequencing of Difficult Templates

Jan Kieleczawa[1,*] *and Erica Mazaika*[2]

[1]*Pfizer Research/Global Biotherapeutics Technologies, Cambridge, Massachusetts 02140, USA; and* [2]*Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA*

In this paper, we have fine-tuned a DNA sequencing protocol suitable for a wide range of difficult templates. The primary goal was to evaluate a number of parameters—such as various dye terminator mixes in the presence or absence of additives, the amount of DNA or primer, and cycling protocols—about the effectiveness of reading through complex regions. We showed that the modification of a published protocol leads to significant (75%) cost reduction without forfeiting quality of the data. In the recommended protocol, we used betaine as a standard additive, but better results can be obtained when betaine and Reagent A are mixed in an equivalent ratio.

KEY WORDS: complex DNA regions, additives, dGTP

INTRODUCTION

Although in the last 5 years, new platforms in next generation sequencing technology have become the focus for sequencing advancement, Sanger sequencing is still far from obsolete. In fact, a DNA Sequencing Research Group (DSRG) survey published in 2007 in the *Journal of Biomolecular Techniques*,[1] consisting of data from 61 laboratories (along with personal communication, J. Kieleczawa, unpublished), indicates that as of now, the number of Sanger reactions performed in many individual core sequencing facilities is still increasing. For example, over the last 3 years, our laboratory has experienced a yearly increase of over 50% in demand for DNA sequencing services. However, the number of capillary instruments decreased dramatically, in some cases, by 90%, in almost all big sequencing centers. Once the reference sequence of an organism of interest is known, one of the next-generation platforms can now be used to sequence the same or similar species relatively quickly without relying on capillary machines. Recently, Roche released a "junior" version of its 454 platform, suited ideally for smaller, individual, next-generation sequencing projects. Coupled, for example, with Life Technology's (Foster City, CA, USA) ABI 3500, an efficient combination may be created to solve many types of sequencing projects.

In general, we believe that the Sanger technology will be viable for many years to come. A relatively straightforward and simple process to obtain good quality and long reads is invaluable for many applications (e.g., gap-closing, resequencing of individual genes), and in many cases, this technology is irreplaceable.

However, certain sequence motifs in DNA templates may interfere with long read lengths, and in these cases, the expert laboratory technician must use one of alternate protocols to yield longer reads through such regions. In our laboratory, the number of reactions requiring enhancement to the standard ABI protocol[2] is 7–10%, at about 10,000–15,000 reactions/year. The protocol we use most often for many types of difficult templates is similar to one from a 2008 DSRG study,[3] which uses two different big dye terminators (BDT) at a specific ratio and in the presence of a zwitterion, betaine. The DSRG study found that the use of full-strength BDT 3.1/dGTP3.0 at a ratio of 3:1 (v/v) in the presence of 1 M betaine (at a cost of $6–7/reaction) will sequence through the widest range of difficult templates.

In this study, we take a more comprehensive approach by studying 16 difficult regions (eight DNA templates, each sequenced in forward and reverse direction around the difficult region). The following variables were evaluated: different BDT 3.1:dGTP3.0 ratios at various dilution strength; various additives; amount of DNA and primers; and cycling conditions. Through optimizing ratios of BDT, using various sequencing additives and modifying cycling conditions, we were able to develop an optimal protocol costing only approximately $1.50/reaction, a savings of $4.50–5.50/reaction, resulting in total savings of $45,000–$55,000/year.

MATERIALS AND METHODS

All materials and methods used in this paper were described extensively in earlier publications.[3–9] The most

*ADDRESS CORRESPONDENCE TO: Jan Kieleczawa, Pfizer Research/Global Biotherapeutics Technologies, 87 Cambridgepark Dr., Cambridge, MA 02140, USA (Phone: 617-665-8318; E-mail: jkieleczawa@wyeth.com; j.kieleczawa@verizon.net).

difficult DNAs were obtained internally from Wyeth scientists during a normal course of submissions and selected based on a region's difficulty. The following cycling conditions were used in all experiments unless otherwise stated: Combine 150 ng DNA; 1 μl 5 μM primer; 10 mM Tris, 0.01 mM EDTA, pH 8.0 (TEsl); and additive (if used), followed by a heat-denaturation step for 5 min at 98°C. If the heat-denaturation step was omitted, it is so noted later in the text or legend. The reaction volume at this stage was set at 7 μl. The enzyme terminator mix (3 μl at different dilution strength) was then added and cycled 40 times: 96°C/10 s, 50°C/5 s, 60°C/2 min. If different cycling parameters were tested, they are indicated later in the text.

The BDT V3.1, dGTPV3.0, Sequence Enhancer Reagent A, and 5× sequence dilution buffer were from Life Technologies. The other additives that we used were betaine and DMSO (Sigma-Aldrich, St. Louis, MO, USA) and GC Melt (Clontech, Mountain View, CA, USA). Following cycle sequencing, unincorporated dye terminators and salts were purified using Performa Dye Terminator Removal V3 96-well plates (EdgeBioSystems, Gaithersburg, MD, USA) and run on an ABI 3730 genetic analyzer using default run parameters.

For each sequencing condition, the quality read lengths (Q>20, 10) were calculated using Sequence Scanner V1.0 software (Applied Biosystems/Life Technologies).

RESULTS AND DISCUSSION

DNA Templates Used in This Study

Characteristics of difficult regions for DNAs used in this study are shown in Table 1. The descriptions were derived using Examine Repeats module in DNASeq DB LIMS developed in the DNA sequencing group at Wyeth/Pfizer.[11] This module is capable of predicting up to seven different sequence motifs, which from a DNA-sequencing standpoint, may be difficult to read through.

Effect of Dye-Terminator Strength and Dye Mixes on the Ability to Sequence through Difficult Regions

The first condition that we examined was the amount (referred to in relation to 8 μl called for in the original protocol[2]; 8 μl undiluted BDT=1×dilution, regardless of the reaction volume) of the BDT 3.1 used alone and/or in conjunction with dGTPV3.0 mixed in different (v/v) ratios. For each difficult region, we used full-strength BDT 3.1, along with 2, 4, and 8× dilutions. When dGTPV3.0 was present, the following ratios (v/v) were tested: 4:1, 3:1, and 2:1, all at dilutions mentioned above. The dilutions were made using 5× sequence dilution buffer, and the final concentration of $MgCl_2$ was kept at 2 mM. In total, we evaluated 16 different dye terminator ratio/dilution combinations that were used in triplicate for each difficult region.

Figure 1A–C shows that reaction conditions with 3:1 and 4:1 ratios of BDT 3.1:dGTPV3.0 in the presence of betaine resulted in the longest read length for the

TABLE 1

Characteristics of Difficult Regions Used in This Study

DNA #	Priming direction	Characteristics of a difficult region
1	F	90% GC over 420 base stretch, 101 base G/C nonrepeat
	R	86% GC over 650 base stretch, 105 base G/C nonrepeat, 7 GCC trinucleotide repeats
2	F	519 base T/C nonrepeat, 2–3 T/C direct repeats of various length (from 12 to 127 bases)
	R	265 base A/G nonrepeat, 2–6 A/G direct repeats of various lengths (from 12 to 51 bases)
3	F	18 and 10 base long homopolymers C separated by TCACCCT
	R	10 and 18 base long homopolymers G separated by AGGGAGA
4	F	24 base long hairpin (60% GC) with $T_m > 95°C$
	R	24 base long hairpin (60% GC) with $T_m > 95°C$
5	F	19 base long homopolymer C (then 200 bases), 19 base hairpin, followed by 41 base A/T nonrepeat containing 15 base hairpin (with 20 bases between hairpin and nonrepeat)
	R	41 base T/A nonrepeat containing 15 base hairpin and 19 base hairpin 20 bases downstream
6	F	Alu repeat and 22 base hairpin with 84 base loop
	R	Alu repeat and 22 base hairpin with 84 base loop
7	F	26 AG dinucleotide repeats
	R	26 CT dinucleotide repeats
8	F	20 TC immediately followed by 23 TG dinucleotide repeats
	R	23 CA followed by 20 GA dinucleotide repeats

F, Forward primer; R, reverse primer.

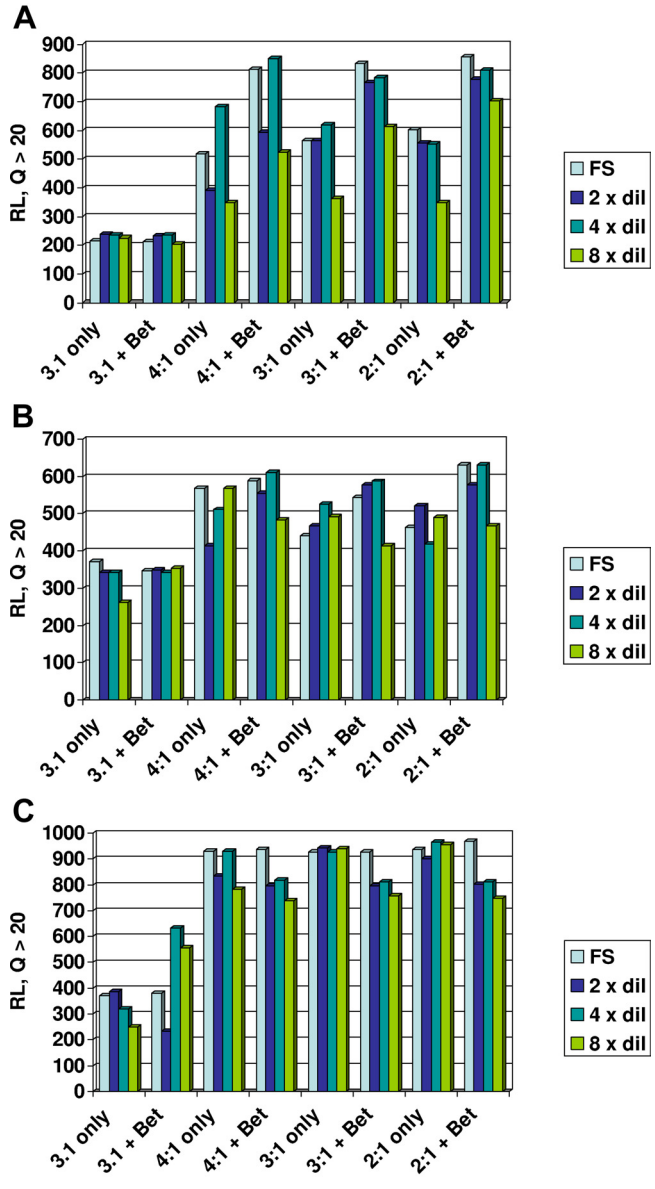

FIGURE 1

Examples of Q ≥ 20 read length (RL) values for different dye dilution strengths in the presence or absence of betaine (Bet). (A) DNA #1 sequenced with forward primer. (B) DNA #5 sequenced with reverse primer. (C) DNA #6 sequenced with forward primer. FS, Full-strength of dye or dye mixes; 2, 4, 8× dil = dye or dye mixes diluted two-, four-, or eightfold, respectively.

majority of difficult regions, and there was no significant decrease in read length when the reactions contained the full-strength 2 or 4× dye dilution. In some cases, read lengths dropped significantly at 8× dilution; hence, the rest of the tests were conducted using 4× dilutions of BigDye 3.1 and dGTPV3.0 at a 3:1 (v/v) ratio. We refer to this combination in the text as a "mix". Table 2A (no betaine) and B (with betaine) summarizes all data from these experiments. Adding betaine to 1 M final concen-

TABLE 2A

Q ≥ 20 Read Length for Forward and Reverse Primers at Different Dye or Dye Mix Dilutions in the Absence of Betaine

BDT Dilution Factor→		Full-strength			2× Diluted				4× Diluted				8× Diluted			
BDT 3.1:→dGTP 3.0 ratio DNA #	BDT3.1 only	4:1	3:1	2:1	BDT3.1 only	4:1	3:1	2:1	BDT3.1 only	4:1	3:1	2:1	BDT3.1 only	4:1	3:1	2:1
1: F	216 ± 10	518 ± 9	563 ± 64	602 ± 23	238 ± 5	391 ± 26	563 ± 40	554 ± 56	235 ± 24	682 ± 77	620 ± 72	552 ± 55	225 ± 14	348 ± 40	362 ± 3	349 ± 8
1: R	551 ± 40	531 ± 12	498 ± 25	505 ± 14	492 ± 10	542 ± 27	558 ± 14	574 ± 11	542 ± 17	586 ± 25	577 ± 13	555 ± 26	504 ± 11	598 ± 12	628 ± 9	614 ± 8
2: F	482 ± 23	555 ± 76	628 ± 27	567 ± 44	479 ± 5	499 ± 56	540 ± 19	554 ± 31	488 ± 44	568 ± 2	572 ± 18	543 ± 35	437 ± 17	502 ± 43	518 ± 14	517 ± 30
2: R	492 ± 5	655 ± 15	634 ± 38	664 ± 4	496 ± 5	562 ± 39	592 ± 17	611 ± 17	501 ± 2	654 ± 25	626 ± 9	639 ± 8	497 ± 1	568 ± 16	581 ± 2	594 ± 19
3: F	119 ± 6	93 ± 3	100 ± 17	84 ± 3	107 ± 8	105 ± 5	97 ± 2	98 ± 2	127 ± 10	95 ± 4	117 ± 10	93 ± 2	128 ± 6	127 ± 6	129 ± 15	136 ± 18
3: R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4: F	467 ± 27	800 ± 25	832 ± 87	913 ± 7	406 ± 2	934 ± 9	889 ± 19	836 ± 22	432 ± 24	968 ± 13	953 ± 21	958 ± 13	269 ± 6	975 ± 15	911 ± 30	898 ± 77
4: R	728 ± 20	853 ± 6	883 ± 11	858 ± 7	514 ± 15	887 ± 30	886 ± 5	888 ± 6	542 ± 57	900 ± 9	903 ± 22	887 ± 29	478 ± 2	909 ± 31	908 ± 30	853 ± 6
5: F	519 ± 14	582 ± 5	548 ± 15	566 ± 19	497 ± 25	605 ± 36	576 ± 30	545 ± 14	527 ± 18	552 ± 16	592 ± 40	594 ± 5	468 ± 5	536 ± 28	537 ± 28	509 ± 9
5: R	371 ± 4	567 ± 3	441 ± 127	463 ± 80	343 ± 8	413 ± 146	468 ± 98	521 ± 94	342 ± 16	511 ± 58	526 ± 73	419 ± 103	261 ± 66	568 ± 10	491 ± 57	490 ± 63
6: F	372 ± 36	931 ± 20	925 ± 27	938 ± 4	387 ± 28	813 ± 99	935 ± 4	947±16	377 ± 23	941 ± 10	920 ± 8	966 ± 8	334 ± 13	756 ± 86	855 ± 6	912 ± 10
6: R	582 ± 12	700 ± 3	653 ± 34	658 ± 10	590 ± 15	616 ± 65	709 ± 9	690 ± 13	602 ± 6	673 ± 9	689 ± 24	659 ± 7	600 ± 6	684 ± 24	724 ± 17	700 ± 10
7: F	684 ± 10	781 ± 19	812 ± 14	878 ± 4	708 ± 28	718 ± 27	710 ± 15	725 ± 11	722 ± 19	786 ± 34	786 ± 19	804 ± 21	677 ± 21	711 ± 69	754 ± 38	816 ± 11
7: R	728 ± 38	398 ± 15	384 ± 12	270 ± 11	495 ± 19	390 ± 21	364 ± 7	340 ± 17	581 ± 52	453 ± 12	405 ± 8	401 ± 12	508 ± 12	404 ± 18	412 ± 15	371 ± 28
8: F	408 ± 30	165 ± 40	152 ± 30	229 ± 47	452 ± 27	252 ± 81	227 ± 65	252 ± 43	413 ± 61	247 ± 31	231 ± 63	209 ± 52	429 ± 65	320 ± 99	255 ± 34	248 ± 43
8: R	712 ± 40	838 ± 23	821 ± 99	826 ± 9	688 ± 63	793 ± 30	817 ± 24	823 ± 7	758 ± 73	803 ± 17	816 ± 21	750 ± 50	707 ± 41	715 ± 55	752 ± 34	811 ± 31

TABLE 2B

Q ≥ 20 Read Length for Forward and Reverse Primers at Different Dye or Dye Mix Dilutions in the Presence of Betaine

BDT Dilution Factor →	Full-strength				2× Diluted				4× Diluted				8× Diluted			
BDT 3.1:→dGTP 3.0 ratio DNA #	BDT3.1 only	4:1	3:1	2:1	BDT3.1 only	4:1	3:1	2:1	BDT3.1 only	4:1	3:1	2:1	BDT3.1 only	4:1	3:1	2:1
1: F	213 ± 16	813 ± 21	831 ± 36	855 ± 7	233 ± 5	593 ± 61	766 ± 17	766 ± 36	234 ± 5	849 ± 38	783 ± 11	810 ± 18	202 ± 34	524 ± 25	614 ± 51	701 ± 21
1: R	603 ± 9	738 ± 19	726 ± 41	738 ± 13	573 ± 10	665 ± 30	742 ± 18	739 ± 23	576 ± 14	832 ± 24	819 ± 9	833 ± 9	567 ± 1	753 ± 34	784 ± 37	811 ± 7
2: F	564 ± 45	623 ± 78	622 ± 74	638 ± 30	540 ± 9	547 ± 16	527 ± 36	493 ± 72	562 ± 69	603 ± 28	577 ± 30	615 ± 9	485 ± 32	520 ± 40	541 ± 23	535 ± 46
2: R	487 ± 4	699 ± 20	708 ± 15	730 ± 7	497 ± 4	619 ± 32	645 ± 40	652 ± 49	497 ± 5	700 ± 18	681 ± 19	699 ± 9	501 ± 1	635 ± 20	657 ± 18	679 ± 19
3: F	207 ± 36	148 ± 16	132 ± 30	101 ± 3	150 ± 9	123 ± 18	114 ± 18	107 ± 12	77 ± 80	113 ± 6	121 ± 6	118 ± 4	167 ± 19	137 ± 11	139 ± 17	122 ± 6
3: R	187 ± 58	862 ± 34	856 ± 5	869 ± 12	0	509 ± 41	529 ± 34	596 ± 40	0	797 ± 35	692 ± 67	760 ± 77	0	369 ± 114	394 ± 11	580 ± 62
4: F	411 ± 14	918 ± 22	902 ± 56	944 ± 9	437 ± 2	917 ± 55	933 ± 51	878 ± 35	429 ± 5	968 ± 16	975 ± 5	936 ± 24	296 ± 28	914 ± 61	979 ± 8	974 ± 3
4: R	903 ± 5	861 ± 8	901 ± 15	922 ± 5	873 ± 13	829 ± 23	862 ± 14	846 ± 9	886 ± 16	909 ± 8	885 ± 31	884 ± 6	620 ± 36	872 ± 46	913 ± 4	871 ± 9
5: F	614 ± 7	835 ± 18	841 ± 8	837 ± 27	619 ± 11	795 ± 21	795 ± 5	802 ± 23	633 ± 15	819 ± 13	813 ± 14	812 ± 15	557 ± 21	739 ± 32	758 ± 24	747 ± 25
5: R	346 ± 3	589 ± 19	543 ± 99	630 ± 91	349 ± 19	554 ± 17	578 ± 98	521 ± 3	342 ± 10	611 ± 32	587 ± 3	631 ± 79	353 ± 2	484 ± 98	413 ± 51	467 ± 96
6: F	381 ± 17	937 ± 27	927 ± 16	967 ± 7	234 ± 19	834 ± 55	941 ± 19	901±66	319 ± 42	930 ± 10	925 ± 23	965 ± 22	248 ± 9	782 ± 37	939 ± 16	954 ± 13
6: R	666 ± 25	893 ± 4	937 ± 11	957 ± 5	603 ± 11	683 ± 8	798 ± 8	929 ± 4	640 ± 7	969 ± 16	954 ± 14	972 ± 17	613 ± 7	754 ± 15	886 ± 31	958 ± 33
7: F	272 ± 23	819 ± 23	824 ± 27	851 ± 1	292 ± 10	658 ± 46	614 ± 42	709 ± 76	293 ± 5	809 ± 43	821 ± 3	887 ± 6	287 ± 19	754 ± 56	678 ± 99	730 ± 54
7: R	859 ± 69	913 ± 23	772 ± 98	805 ± 48	855 ± 25	774 ± 50	860 ± 5	676 ± 57	869 ± 28	883 ± 16	789 ± 67	747 ± 66	787 ± 72	863 ± 6	711 ± 97	651 ± 57
8: F	493 ± 48	244 ± 26	300 ± 42	314 ± 56	290 ± 47	281 ± 57	304 ± 14	341 ± 23	202 ± 46	236 ± 53	272 ± 48	294 ± 67	147 ± 4	252 ± 34	324 ± 30	283 ± 30
8: R	527 ± 8	790 ± 17	767 ± 36	828 ± 4	518 ± 7	782 ± 73	772 ± 61	746 ± 34	528 ± 10	762 ± 52	770 ± 16	737 ± 56	480 ± 82	702 ± 49	783 ± 23	730 ± 33

TABLE 3

Effect of Various Additives on Q ≥ 20 Read Length for Three Most Difficult DNA Regions

DNA # and Primer Direction→	1-F		1-R		2-F		2-R		4-F		4-R	
± Heat denaturation (± HD)→	− HD	+ HD	− HD	+ HD	− HD	+ HD	− HD	+ HD	− HD	+ HD	− HD	+ HD
BDT3.1 only (1/4 dilution)	65 ± 8	251 ± 3	252 ± 45	548 ± 3	469 ± 45	450 ± 30	495 ± 3	501 ± 5	177 ± 4	323 ± 31	429 ± 8	507 ± 3
Mix only (1/4 dilution)	79 ± 1	452 ± 25	250 ± 2	661 ± 44	584 ± 3	561 ± 9	615 ± 55	648 ± 32	910 ± 47	969 ± 11	935 ± 11	960 ± 4
Mix + 2 µl betaine	687 ± 148	916 ± 22	321 ± 81	960 ± 38	680 ± 13	571 ± 111	719 ± 17	757 ± 27	852 ± 74	964 ± 22	879 ± 98	965 ± 19
Mix + 3 µl betaine	933 ± 8	964 ± 17	893 ± 21	985 ± 5	490 ± 242	571 ± 58	642 ± 15	704 ± 26	356 ± 175	761 ± 74	456 ± 41	761 ± 74
Mix + 2 µl Reagent A	875 ± 42	927 ± 15	930 ± 14	982 ± 11	688 ± 57	573 ± 27	658 ± 30	743 ± 20	907 ± 31	958 ± 6	907 ± 31	958 ± 6
Mix + 3 µl Reagent A	890 ± 80	942 ± 13	909 ± 48	974 ± 20	479 ± 235	571 ± 34	512 ± 23	526 ± 3	263 ± 143	795 ± 46	942 ± 1	864 ± 101
Mix + 1 µl betaine + 1 µl Reagent A	869 ± 16	964 ± 10	947 ± 34	1000 ± 8	697 ± 43	582 ± 15	738 ± 22	779 ± 15	926 ± 15	974 ± 6	948 ± 6	965 ± 7
Mix + DMSO (5% final)	465 ± 147	908 ± 19	475 ± 107	686 ± 86	522 ± 22	476 ± 28	657 ± 31	654 ± 29	949 ± 28	949 ± 7	956 ± 16	938 ± 4
Mix + 2 µl betaine + DMSO	965 ± 11	969 ± 19	876 ± 12	937 ± 7	517 ± 117	451 ± 182	651 ± 27	634 ± 18	543 ± 167	861 ± 36	546 ± 167	861 ± 36
Mix + 2 µl Reagent A + DMSO	940 ± 32	982 ± 22	885 ± 21	910 ± 16	435 ± 159	596 ± 33	525 ± 7	528 ± 2	379 ± 180	796 ± 61	389 ± 282	794 ± 77
Mix + 2 µl betaine/Reagent A + DMSO	192 ± 45	326 ± 31	579 ± 28	690 ± 40	0	0	86 ± 33	0	60 ± 50	82 ± 0	406 ± 0	252 ± 240
Mix + 2 µl GC Melt	865 ± 29	908 ± 10	815 ± 71	936 ± 42	725 ± 7	648 ± 17	703 ± 59	758 ± 24	870 ± 3	935 ± 1	953 ± 23	940 ± 13
Mix + 2 µl betaine + 2 µl GC Melt	422 ± 81	818 ± 150	737 ± 55	856 ± 50	113 ± 18	326 ± 11	457 ± 60	542 ± 31	171 ± 3	249 ± 23	589 ± 108	699 ± 31
Mix + 2 µl Reagent A + 2 µl GC Melt	623 ± 134	754 ± 107	867 ± 29	952 ± 25	136 ± 2	178 ± 102	163 ± 3	492 ± 3	143 ± 14	248 ± 24	329 ± 318	390 ± 234
Mix + 2 µl GC Melt + DMSO	972 ± 9	987 ± 20	811 ± 60	924 ± 32	423 ± 18	504 ± 31	570 ± 5	570 ± 13	523 ± 29	767 ± 51	563 ± 19	797 ± 39
Mix + 2 µl betaine/Reagent A + 2 µl GC Melt + DMSO	56 ± 0	147 ± 2	117 ± 47	435 ± 80	0	0	0	0	0	0	0	0

+ HD or − HD indicates whether the 5-min heat-denaturation step was (+) or was not (−) part of the sequencing protocol. Q ≥ 20 values highlighted in blue represent the longest reads in a given category. Q ≥ 20 values highlighted in red are within 10% of the best values.

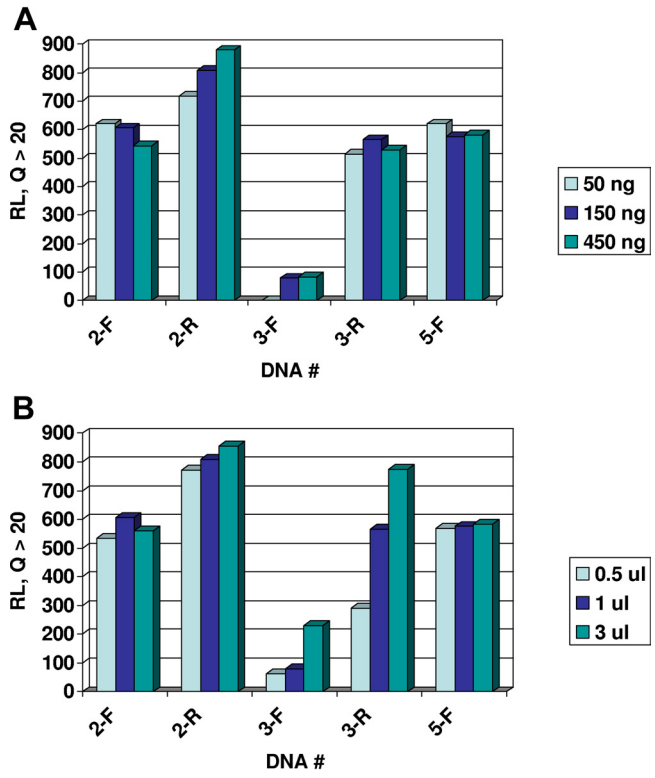

FIGURE 2

Effect of amount of DNA (A) or primer (B) on Q ≥ 20 read-length values. Four different templates were sequenced at three levels of DNA or primer. Template #3 was sequenced in both directions, as this is the most difficult template in this series.

tration almost always improved read length and data quality.

Effect of Additives

To enhance further the effectiveness of the above sequencing mix, we tested various commercially available additives on the quality of reads for three most difficult regions from our DNA panel. Adding 1 μl Reagent A and 1 μl betaine to a mix produced the best quality data. Using Reagent A or GC Melt only was also effective. Routinely, however, we recommend using betaine as a result of its lower cost. Table 3 summarizes all data.

Effect of the Amount of DNA and Primer

Anecdotal information heard at various meetings suggested that the amount of DNA and/or primer may have some influence on the ability to read through some difficult regions. To evaluate such a possibility, we selected three DNA templates and varied the amount of DNA from 50 to 450 ng/reaction (Fig. 2A) and then set the amount of DNA at 150 ng/reaction and varied the amount of 5 μM primer from 0.5 to 3 μl (Fig. 2B). In both cases, there was no significant effect of the amount of DNA or primer on the read length. The data for DNA #3 (R primer) were misleading and were a result of the inability of Phred to call Q ≥ 20 values accurately for difficult DNAs.[3] Visual inspection of these data indicated that in each case, only 90–100 bases can be called accurately.

Testing Various Cycling Conditions

Six different cycling regimes were evaluated on six different DNA templates to come up with the most efficient protocol suitable for the widest array of difficult regions. The details/descriptions of these protocols are below:

Protocol #1. Combine DNA, primer, additive, and TEsl. Total volume at this point was 7 μl. Heat-denature this mix for 5 min at 98°C. Cool down on ice. Add 3 μl dye terminator (or dye mix), and cycle: [(96°C/10 s) (50°C/5 s) (60°C/2 min)] × 40.[4] The additive can be present during the heat-denaturation step or can be added as a part of dye mix.[4] For convenience reasons, we prepared dye-additive mix and stored it at −20°C for a few weeks without loss of its effectiveness (data not shown).

Protocol #2. Combine all components of sequencing reactions (10 μl), and hold this mix for 1 min at 96°C. Then cycle: [(96°C/10 s) (50°C/5 s) (60°C/4 min)] × 30. This is essentially the standard ABI-like sequencing protocol.[2]

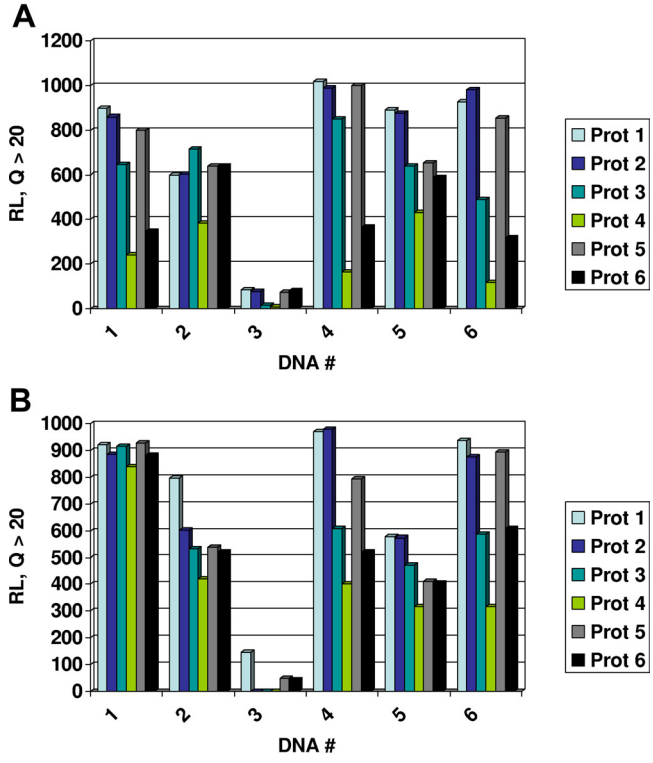

FIGURE 3

Testing various cycling conditions on six different DNAs with forward (A) or reverse (B) primers.

Protocol #3. Similar to Protocol #1, but the extension is set at 65°C (not 60°C).

Protocol #4. Similar to Protocol #1, but the extension is set at 70°C (not 60°C).

Protocol #5. Similar to Protocol #3, but the transition time between annealing temperature (50°C) and extension temperature (65°C) was slowed down to 30 s as opposed to a typical time of approximately 5 s.

Protocol #6. Similar to Protocol #4, but the transition time between annealing temperature (50°C) and extension temperature (70°C) was slowed down to 30 s as opposed to a typical time of approximately 5 s.

Figures 3A (forward primer) and B (reverse primer) show Q ≥ 20 data for all tested DNA templates. In eight of 12 cases, Protocol #1 gave reads that are slightly longer, and the visual inspection in Sequencher indicated that they have overall better quality. Only in two cases did Protocol #2 give longer reads compared with Protocol #1. For the remaining two cases, data generated using either protocol gave data of similar read length. All other protocols produced data that gave significantly shorter than Q ≥ 20 reads, with the possible exception of Protocol #5, which gave comparable read length; however, the quality of data was generally worse, as illustrated in Figure 4A and B. In addition, this protocol added approximately 20 min more time to the cycling regime, and we do not recommend using it.

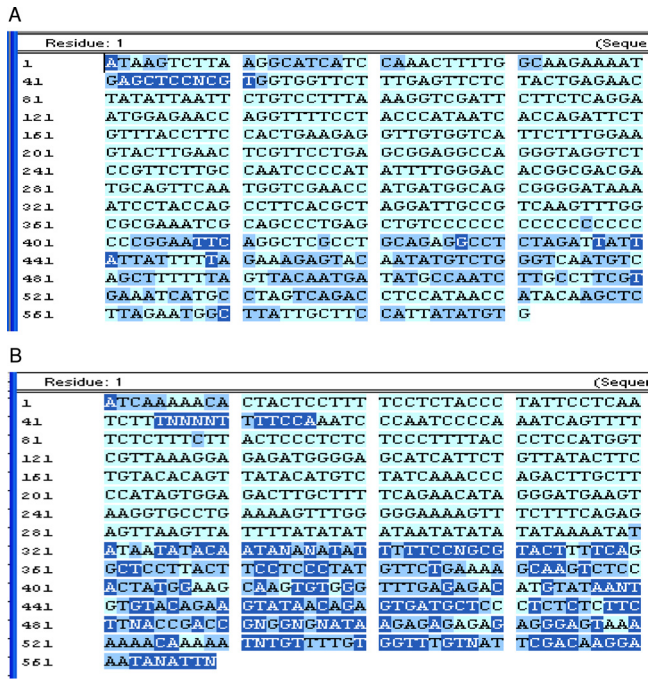

FIGURE 4

Comparison of data quality using cycling Protocol #1 (A) or Protocol #5 (B). Note that the amount of dark-blue bases is higher (and starts earlier) in the example sequenced with Protocol #5. In Sequencher's notation, the light-blue color indicates bases with Q ≥ 20. Darker blue colors of bases indicate their poorer quality.

Summary

In this study, we have tested an extensive range of parameters to develop the most efficient yet cost-effective protocol for sequencing a wide range of difficult templates. The most optimal protocol included a mixture of BDT 3.1 and dGTPV3.0 at a v/v ratio of 3:1 and in the presence of 1 μl betaine and 1 μl Reagent A. However, taking into account the cost factor/reaction ($0.02 for betaine only and $1.01 for betaine/Reagent A mix), we recommend using betaine only.

ACKNOWLEDGMENTS

We thank many Wyeth scientists who provided us with DNA templates with a variety of difficult regions. We also recognize our appreciation to our colleagues at the DNA Sequencing Group in Cambridge for valuable suggestions and discussions. Finally, we express our gratitude to Global Biological Technologies management for creating an environment where such studies are encouraged and possible.

REFERENCES

1. Needleman D, Adam D, Detwiler M, et al. DNA Sequencing Research Group: 2006 General Survey of DNA Sequencing Facilities. *J Biomol Tech* 2007;18:113–119.
2. ABI PRISM® BigDye™ Terminator v3.1 Cycle Sequencing Kit. Protocol. Foster City, CA, USA: Applied Biosystems, 2002;Part number 4337035 Rev. A.
3. Kieleczawa J, Adam D, Bintzler D, et al. Identification of optimal protocols for sequencing difficult templates: the results of the ABRF DSRG Difficult Template Study 2008. *J Biomol Tech* 2009;20:116–127.
4. Kieleczawa J. Sequencing of difficult DNA templates. In Kieleczawa J (ed): *DNA Sequencing: Optimizing the Process and Analysis*, Sudbury, MA, USA: Jones and Bartlett, 2005;27–34.
5. Kieleczawa J. Simple modifications of the standard DNA sequencing protocol allow for sequencing through siRNA hairpins and other repeats. *J Biomol Tech* 2005;16:220–223.
6. Kieleczawa J. Controlled heat-denaturation of DNA plasmids. In Kieleczawa J (ed): *DNA Sequencing: Optimizing the Process and Analysis.* Sudbury, MA, USA: Jones and Bartlett, 2005;1–10.
7. Yie Y, Wei Z, Tien P. A simplified and reliable protocol for plasmid DNA sequencing: Fast miniprep and denaturation. *Nucleic Acids Res* 1993;21:361.
8. Kieleczawa J, Wu P. Prolonged storage of plasmid DNAs under different conditions: effects on plasmid integrity, spectra characteristics, and DNA sequence quality. In Kieleczawa J (ed): *DNA Sequencing II: Optimizing the Preparation and Clean Up*, Sudbury, MA, USA: Jones and Bartlett, 2005;99–115.
9. Kieleczawa J, Li T, Wu, P. Preparation of difficult DNA templates using seven different commercial methods. In Kieleczawa J (ed): *DNA Sequencing II: Optimizing the Preparation and Clean Up*, Sudbury, MA, USA: Jones and Bartlett, 2006;1–14.
10. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 1998;8:175–185.
11. Kieleczawa J, Koffman D, Lakshmanan B, Kitzmiller A. Bioinformatics tools to aide sequencing of difficult templates In Kieleczawa J (ed): *DNA Sequencing III: Dealing with Difficult Templates*, Sudbury, MA, USA, Jones and Bartlett, 2008;163–177.