IN SILICO SEQUENCE OPTIMIZATION FOR THE REPRODUCIBLE GENERATION OF DNA STRUCTURES

by

Michael D. Tobiason

A dissertation

submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy in Material Science and Engineering

Boise State University

December 2019

© 2019

Michael D. Tobiason

ALL RIGHTS RESERVED

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the dissertation submitted by

Michael D. Tobiason

Dissertation Title: *In Silico* Sequence Optimization for the Reproducible Generation of DNA Structures

Date of Final Oral Examination: 15 October 2019

The following individuals read and discussed the dissertation submitted by student Michael D. Tobiason, and they evaluated the student's presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

William L. Hughes, Ph.D.	Chair, Supervisory Committee
Bernard Yurke, Ph.D.	Member, Supervisory Committee
Jeunghoon Lee, Ph.D.	Member, Supervisory Committee
James Alexander Liddle, Ph.D.	Member, Supervisory Committee
Igor L. Medintz, Ph.D.	Member, Supervisory Committee

The final reading approval of the dissertation was granted by William L. Hughes, Ph.D., Chair of the Supervisory Committee. The dissertation was approved by the Graduate College.

DEDICATION

I dedicate this dissertation to three groups of people. First and foremost, my unbelievably supportive family. Colleen, Bernie, Steve, Andy, Carlie, Larkin, and Jon; without your support and encouragement I would never have made it this far. Second, I owe much of who I am to the countless coaches and teachers who have mentored me throughout the years. I owe a special thank you to Tom Shanahan, Patrick McCurry, Larry Neznanski, Guy Hudson, Jim Dull, Katie Devine, and Will Hughes; Hopefully all those second chances you gave me have officially started paying off. Third and finally, thank you to all those friends and family who have chosen to lift the sword instead of the pen. Pat Ellison, Sarah Belmont, and many more; thank you for fighting so that I have the freedom of thought which I cherish so much.

ACKNOWLEDGMENTS

Thank you first and foremost to my dissertation committee for the time and effort they have spent shaping this project: Dr Hughes, Dr. Yurke, Dr. Lee, Dr. Liddle, and Dr. Medintz. Your feedback has been fundamental to improving both the science and my approach to the science. Special thanks to Bernie for sitting down with me in September 2014 and teaching me to code in Java. I never expected that the relatively simple program we created would evolve into this PhD dissertation. Another special thank you to Will for his limitless patience with me over the years. I have learned more from you than I will ever be able to explain or put into words.

Thank you to all the faculty and staff at Boise State who I have had the pleasure to work with. Special thanks to Natalya Hallstrom, Elton Graugnard, Jennifer Padilla, Reza Zadegen, Wan Kuang, Bill Knowlton, Paul Davis, and Donald Kellis.

Finally, thank you to all my fellow graduate students who made day-to-day life at Boise State such a pleasure. Special thanks to Brett Ward, Brittany Canon, Sara Goltry, and Sadao Takabayashi; I could never have survived graduate school without you.

ABSTRACT

Biologically, deoxyribonucleic acid (DNA) molecules have been used for information storage for more than 3 billion years.[1] Today, modern synthesis tools have made it possible to use synthetic DNA molecules as a material for engineering nanoscale structures. These self-assembling structures are capable of both resolutions as fine as 4 angstroms and executing programed dynamic behavior.[2,3] Numerous approaches for creating structures from DNA have been proposed and validated, however it remains commonplace for engineered systems to exhibit unexpected behaviors such as low formation yields, poor performance, or total failure. It is plausible that at least some of these behaviors arise due to the formation of non-target structures, but how to quantify and avoid these interfering structures remains a critical question.

To evaluate the impacts of non-target structures on system behavior, three co-dependent scientific developments were necessary. First, three new optimization criteria for quantifying system quality were proposed and studied. This led to the discovery that relatively small intramolecular structures lead to surprisingly large deviations in system behavior such as reaction kinetics. Second, a new heuristic algorithm for generating high quality systems was developed. This algorithm enabled the experimental characterization of newly generated systems, thus validating the optimization criteria and confirming the finding that almost all kinetic variation can be explained by non-target intramolecular structures. Finally, these studies necessitated the creation of two new software tools; one for analyzing existing DNA systems (the “Device Profiler” software) and another for

generating fit DNA systems (the “Sequence Evolver” software). In order to enable these tools to handle the size and complexity of state-of-the-art systems, it was necessary to invent efficient software implementations of the metrics and algorithm. The performance of the software was benchmarked against several alternative tools in use by the DNA nanotechnology community, with the results indicating a marked improvement in system quality over current state-of-the-art methods. Ultimately, the new optimization criteria, heuristic algorithm, and software cooperatively enabled an improved method for generating DNA systems with kinetically uniform behaviors.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3’-	The 3’ end of a DNA molecule’s phosphate chain
3D	Three Dimensional
5’-	The 5’ end of a DNA molecule’s phosphate chain
A	Adenosine
Approx.	Approximately
AFM	Atomic Force Microscopy
B-DNA	Double helical structure adopted by two complementary DNA strands. B is a historical reference to the name of the second of Franklin and Gosling’s X-ray powder photographs.
bp	Base Pair
BHQ2	Black Hole Quencher 2. A proprietary dye-quenching molecule
BSU	Boise State University
C	Cytosine
CPL	Cycles-Per-Lineage
DevPro	Device Profiler (custom written software)
DNA	Deoxyribonucleic acid (class of biomolecules)
eq.	Equation
e.g.	Latin “exempli gratia” (literally meaning “for example”)
et al.	Latin “et alia” (literally meaning “and others”)
G	Guanine
GPC	Generations-Per-Cycle
HPC	High Performance Computing

ml	milliliter
NFS	Network Fitness Score
NDPM	Number-of-Daughters-Per-Mother
NL	Number-of-Lineages
NMPC	Number-of-Mothers-Per-Cycle
nM	nano-Molar
nm	nanometer
№	Number of
SeqEvo	Sequence Evolver (custom written software)
SFS	Strand Fitness Score
SS	Sequence Symmetry (property for describing system quality)
SSM	Sequence Symmetry Minimization (sequence optimization method)
T	Thymine
TE	Tris-EDTA (chemical buffer)
TEM	Transmission Electron Microscopy
TFS	Total Fitness Score

CHAPTER ONE: INTRODUCTION

Section 1.1 – Motivation

The allure of self-assembly has long fascinated scientists. To understand why, one needs to look no further than a mirror. Each adult human starts as little more than a single cell. This cell begins a replication cascade leading to approximately 30 trillion cells and a host of interconnected non-cellular structures.[4,5] These structures have feature sizes ranging from Angstroms (e.g. chemicals such as DNA) to meters (e.g. extremities such as legs) and are produced with remarkable precision. In a sense, we do not create new humans; we instead create a single cell, which then autonomously fabricates a new human. In today's terms this may be called a bottom-up self-assembly process. As such, we exist as a proof-by-example that systems of incredible complexity can be created using such techniques. In comparison, state-of-the-art synthetic self-assembled structures remain somewhat trivial. However, even the relatively simple structures already achievable are showing potential to revolutionize society in applications such as medical diagnostics and information storage.[6-12]

It turns out that synthetic DNA is a great medium for creating self-assembling structures. This is nicely demonstrated by the university logos produced by the Nanoscale Materials and Devices group here at Boise State (Figure 1-1). These logos are typical of self-assembled structures in the sense that they spontaneously form when the proper reactants are combined. This makes it possible to fabricate a large number of structures simultaneously, with the resulting structures in this example being composed purely of

DNA. In biological self-assembly, the information stored in DNA directs the assembly of structures form a variety of larger materials. As such, it is unlikely that systems created solely from DNA will ever rival the complexity of biological systems. However, advancements in our ability to understand and control DNA structures may yield both technological advancements in the near term and remarkable technologies in the long term.

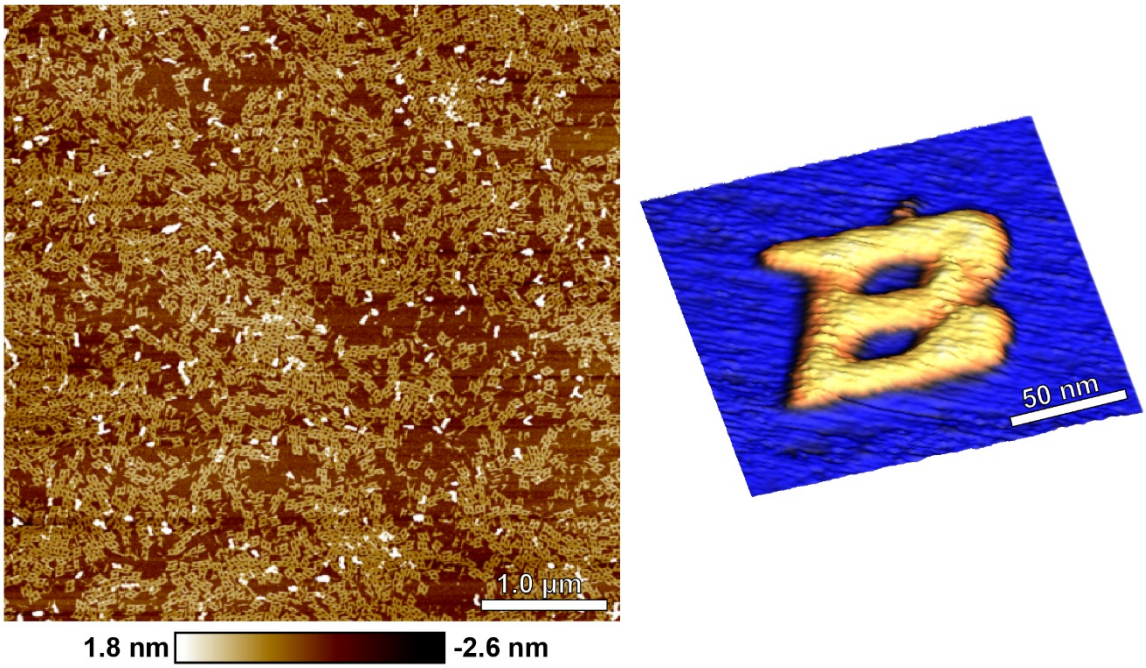


Figure 1-1 The Boise State logo self-assembled using synthetic DNA strands. (left) Atomic force microscopy image showing numerous structures. (right) Isolated and processed image of a single structure. Structure design by Kelly Schutt. Images captured and processed by Brett Ward and Elton Graugnard.

This dissertation will detail an important advancement toward these goals. I will begin by explaining the basics of DNA structures, and how these have been used to create self-assembling shapes. Next, I will present three criteria we have developed for quantifying the quality of a given DNA system. After this, I will explain how we developed an evolution-inspired algorithm for robustly identifying high-quality

structures. Finally, I will explain how we combined our criteria and algorithm to create a software package for automating the design process. I will conclude by discussing the context of these projects, and how I anticipate they may impact the scientific community.

Section 1.2 – The Structure of a DNA Molecule

Chemically, DNA molecules are composed of six sub-structures: deoxyribose, phosphoric acid, and four nucleobases (Figure 1-2, Left). The four nucleobases are Adenine, Cytosine, Guanine, and Thymine (abbreviated A, C, G, and T, respectively). A deoxyribose can be covalently bonded both with phosphoric acid and one of the four nucleobases to create a structure termed a nucleotide. The resulting structure has two exposed OH groups which are referred to as the 5' site (attached to the phosphorus) and the 3' site (part of the deoxyribose). Nucleotides are linked by the covalent bonding of one nucleotide's 5' site with the 3' site of the other. This enables the creation of linear chains of polynucleotides which are commonly known as DNA molecules or DNA strands (Figure 1-2, right). Consequently, the chemical structure of a given DNA strand can be fully described by listing the sequence of its bases starting from the 5' or 3' end of the structure (for example, 5'-GCAT-3' in Figure 1-2 right). This structural motif implies that for DNA molecules containing L bases, there are 4^L possible base sequences. For example, the sequence presented in Figure 1-2 contains only four bases; one of 256 possible four-base sequences. Similarly, for strands 8 bases in length there are 4^8 (approx. 6×10^4) possible sequences, and for strands 16 bases long there are 4^{16} (approx. 4×10^9) possible sequences. Short DNA oligomers (sequence length < approx. 200 bases) of arbitrary sequence can be synthesized directly from chemical precursors and are commercially available.

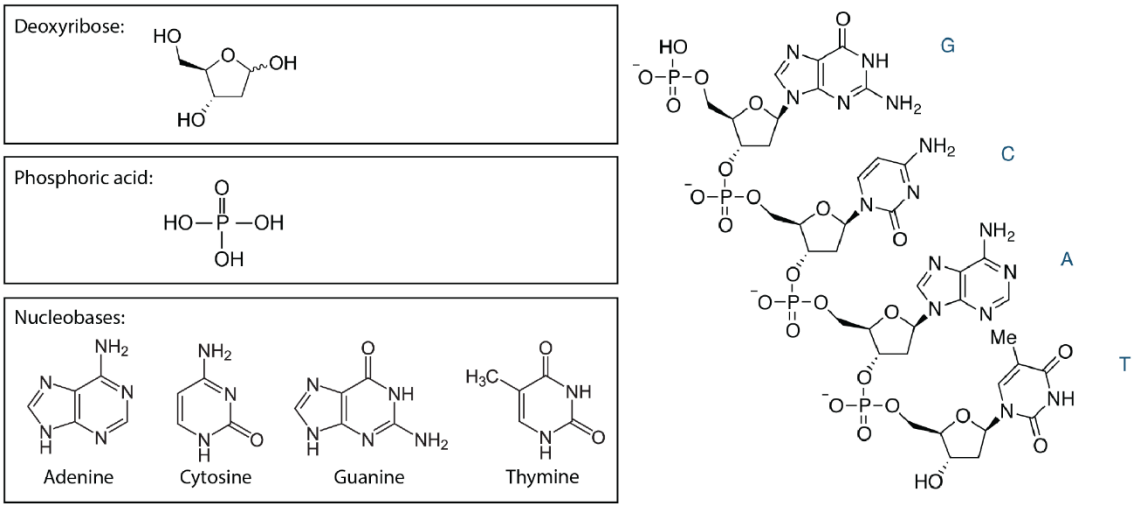


Figure 1-2 Chemical structure of DNA molecules. (Left) Chemical structure of the four sub-structures of which DNA molecules are composed: deoxyribose, phosphoric acid, and the four nucleobases. (Right) Chemical structure of an example DNA molecule containing the sequence 5'-GCAT-3'. Images adapted from the atdbio website (www.atdbio.com).

Section 1.3 – The B-DNA Structure

Circa 1953 it was recognized that certain DNA sequences adopt structures now known formally as B-DNA and colloquially as the DNA double-helix.[13,14] The B-DNA structure arises from the pairing of complementary bases (A with T or C with G) arranged on oppositely-oriented phosphate backbones (e.g. 5'-ACTG-3' and 5'-CAGT-3' in Figure 1-3a). The binding of a single complementary base is relatively unstable, but structural stability increases for stretches of complementary bases. Such stretches of complementary bases are commonly referred to as either domains or simply complements.[15] The size of a complement is often discussed using its length; i.e. the number of complementary bases within the structure with units of either base-pairs or simply bases.

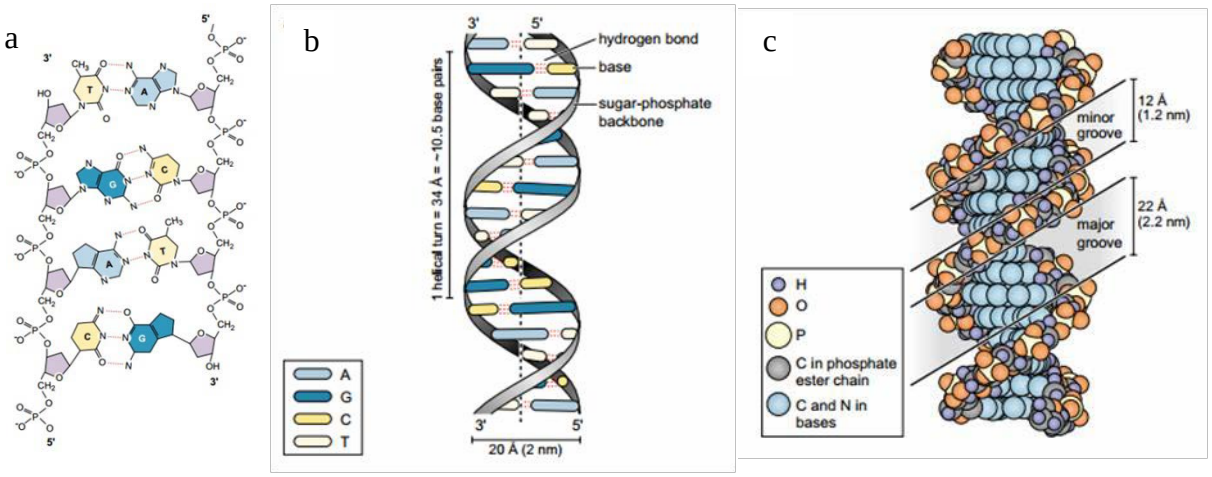


Figure 1-3 Three illustrations of B-DNA structures. (a) Chemical structure of two complementary DNA 4-mers. (b) A cartoon depicting two interacting DNA 16-mers in the double-helical B-DNA structure. (c) An atomic representation of a 17 base-pair B-DNA structure. Each schematic is adapted from Molecular Biology of the Gene by J. D. Watson.[16]

Spatially, B-DNA structures adopt the iconic double-helical shape which is emblematic of DNA (Figure 1-3b,c). This structure has a diameter of approximately 2 nm when hydrated and a righthanded orientation. B-DNA has a repeat unit of ~10.5 base-pairs per helical turn and includes both major and minor grooves (Figure 1-3c). These grooves are approximately 2.2 nm and 1.2 nm in width respectively. DNA molecules have been demonstrated to form a host of other structures including triplex and quadruplex structures; however B-DNA is expected to be the dominant structure under typical experimental conditions.[17] In fact, certain DNA structures form via the same complementary sequences however are not B-DNA. Therefore, when discussing complementary DNA sequences, it is more appropriate to describe their propensity to form B-DNA type structures instead of specifically B-DNA itself.

Section 1.4 – Creating Shapes and Assemblies Using B-DNA

Based on the specificity of B-DNA binding it is possible to rationally design sets of DNA strands – referred to here as DNA systems – which form target structures. Many

methods for rationally designing B-DNA based structures have been proposed.[2,18-31] This spectrum of design methods includes both approaches for designing systems which adapt a target shape (including DNA Origami,[22] DNA-Bricks,[30] and DNA Crystals[18]) and approaches for designing systems which execute programmed chemical dynamics (such as entropy-driven substrates,[25] seesaw gates,[29] and catalyzed hairpin reactions[26,32]). For example, consider the DNA-bricks technique illustrated in Figure 1-4 below.[30,33] In this method, many small synthetic DNA strands are used as nanoscale bricks to assemble a 3D target structure. This is accomplished by first drawing the target shape on a 3D canvas composed of cubes (Figure 1-4a). DNA strands composed of four binding- domains are then created and mapped onto the target structure such that each cube is replaced by a B-DNA structure connecting two strands (Figure 1-4b). DNA sequences for each domain are generated and the corresponding strands are synthesized. Finally, the strands are combined into a single sample and the target structure self-assembles. The structures can then be experimentally characterized using techniques such as Atomic Force Microscopy (AFM) or Transmission Electron Microscopy (TEM) (Figure 1-4c).

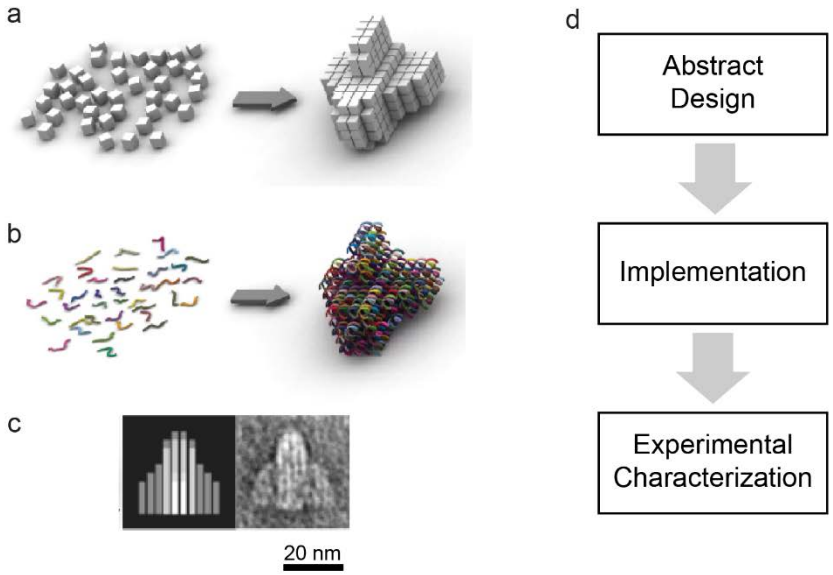


Figure 1-4 Typical state-of-the-art design process illustrated using the DNA Bricks architecture. (a) In the DNA-Bricks architecture target structures are rendered on an abstract 3D canvas where B-DNA duplexes are represented by cubic voxels. (b) DNA strands with sequences implementing the target structure are generated. (c) Strands are chemically synthesized, assembled into the target structure, and are experimentally characterized (TEM Image). (d) Diagram illustrating three key steps in the design process. Images a-c adapted from Ke *et al.*[30]

Section 1.5 – State of the Art Design Procedures

Most state-of-the-art design methods follow a workflow similar to that of DNA-bricks (Figure 1-4d). First, target structures are architected using some level of abstraction. Once the system is designed at this abstract level, strand-sequences are generated which will implement the target structures. This leads to a hierarchical relationship where there are (typically) many possible sequence-sets implementing a single design. In practice this type of design method guarantees that the resulting sequence-sets will contain the target structures. However, they are oblivious to possible alternative structures unless an additional analysis step is introduced. To rectify this, Dirks *et al.* postulated that successful sequence generation methods must optimize both for target structure stability and against non-target structure stability.[34]

Section 1.6 – Non-Target Structures

To better understand the issue of non-target structures and the importance of sequence optimization, consider the model system presented in Figure 1-5 below. First, the goal of the system is established as creating two strands which will form a single B-DNA structure (Figure 1-5a). Towards this goal, we can create a domain-level design such as that in Figure 1-5b. In this design, one strand is composed of the alpha domain and the other the complement of the alpha domain (underlined alpha). This design can then be implemented by assigning specific sequences to the alpha domain such as those in Figure 1-5c. For the example implementation, we have chosen the six-base sequence 5-AATTCG-3 for alpha and this implies the sequences for the remainder of the system. As such, this system is one of 4^6 = 4,096 possible sequence-sets implementing the design.

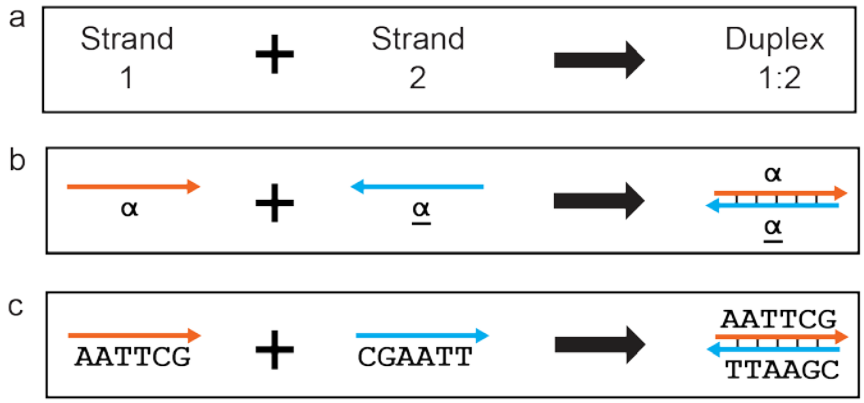


Figure 1-5 A simple model system described at three levels of decreasing abstraction. (a) Strand-level abstraction. (b) Domain-level abstraction. (c) Sequence-level abstraction.

While all systems implementing a design will necessarily contain the target structures, many also contain non-target structures. For example, consider again our simple model system. The generated sequence-set contains only two strands composed of six bases each (Figure 1-6a). By iterating through all possible base pairings, it is possible to exhaustively calculate the B-DNA type structures which the system may form. By

discarding any structure which exists as part of a larger structure, the list of all unrelated or "unique" B-DNA structures can be reported (Figure 1-6b). For this example, we see that there are five total intermolecular structures which are unique. Of these five structures, only the AATTCG/CGAATT structure is a target structure, making the remaining four non-target structures. A subset of structures may form from base pairing of a strand with itself and are termed intramolecular B-DNA type structures. The list of these intramolecular structures can be similarly calculated (Figure 1-6c).

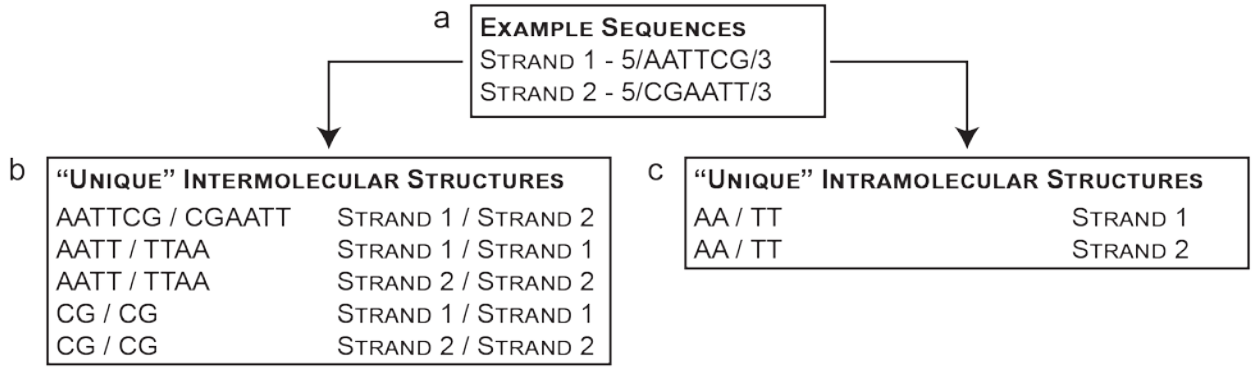


Figure 1-6 B-DNA type structures which the simple model system may form. (a) The example system is composed of two strands named Strand-1 and Strand-2. (b) The list of all unrelated or "unique" intermolecular B-DNA type structures this system may form. (c) The list of all intramolecular B-DNA type structures this system may form.

Section 1.7 – Defining Structural and Dynamic Behavior

If a DNA system is defined as a set of strands with declared sequence (such as those in Figure 1-6a), then a system's structural behavior can be thought of as the set of structures which the strands form. In this sense, good behavior is the formation of target structures, and bad behavior is presence of any other structures. These definitions can be expanded to include a system's dynamic behavior, where the rate of a given transition is either good or bad depending on its proximity to a target rate.

There are a host of issues which are known to lead to variations in structural or dynamic behavior including: blunt end stacking,[35] duplex breathing,[29] thermodynamically favorable non-target structures,[36] thermodynamically unfavorable non-target structures,[37] and structure dimerazation.[38] In theory there may be a nice distinction between structural and dynamic behavior, but it is not uncommon for an issue impacting one to also impact the other. As an example, consider again the simple model system presented in Figure 1-5 above. An experimental sample for this system typically begins with single-stranded reactants which proceed to form B-DNA duplexes. Since the target structure is the B-DNA duplex, this system begins with a high degree of bad structural behavior and makes a transition towards good behavior as time progresses. A slow enough reaction rate effectively locks the system into bad structural behavior, illustrating a kinetic issue leading to bad structural behavior. This type of dependency makes it difficult to parse out which issues are dynamic in nature and which are structural. Furthermore, it highlights that in order to reliably produce good behavior within the time scale of interest, it is necessary to simultaneously address both types of issues.

Section 1.8 – Sequence Symmetry Minimization

The first known method for eliminating non-target structures was the Sequence Symmetry Minimization (SSM) technique pioneered by Seeman circa 1982.[18] Briefly described, SSM is a method for algorithmically generating strand sequences such that the resulting set of strands contains no non-target structures larger than a pre-established threshold. As a result, system quality in SSM is quantified by the size of the largest non-target structure the system may form. This value is referred to here as the Sequence-Symmetry (SS) criterion and is typically described with units of base-pairs (bp). For

example, the system presented in Figure 1-6 would have a SS value of 4 bp due to the presence of 4-base intermolecular structures. This definition of the SS criteria is subtly different from Seeman's original definition since the original definition was tied to the size of the building block and thus typically equal to one bp less than our definition. We have made this distinction for two reasons: First, this definition enables us to analyze existing systems using the criteria, and second this definition is a more intuitive measure of system quality. To date, the SS criterion remains a common metric for discussing the quality of DNA systems.[39] While the scope of the SSM technique was originally limited to the design of nucleic acid junctions and lattices,[18,40] the work has been expanded to enable the creation of other structures.[40] In addition, the success of the SSM technique has inspired the development of additional strategies and tools such as the DNA Sequence Generator (abbreviated DSG) and the Exhaustive Generation of Nucleic Acid Sequences (abbreviated EGNAS) software tools.[41,42] These tools have been shown to generate sequences faster and/or for an expanded class of systems relative to traditional SSM, but still fundamentally rely on the SS criterion for quantifying system quality.

To date, there are currently no fewer than twelve computer programs available for the design and implementation of DNA systems, most of which apply some form of *in silico* analysis to guide the sequence generation process.[15,18,22,41,43-51] These programs evaluate system quality using one of three types of criteria: the SS criterion,[40-42,48] simulated thermodynamic properties,[43,50,52] or other individually developed fitness functions.[15,36,51] Currently, the most precise methods for predicting dynamic behavior rely on individually developed fitness functions, with the most accurate method being the "6-factor" model developed by Zhang *et al.*.[36] This model was generated by training a

prediction algorithm using experimentally measured reaction rates, and the resulting model predicts rates within a factor of 3 for 91% of sequences. Based on this approach, the model is expected to accurately predict reaction rates only for 36 bp duplex-formation reactions under the specific experimental conditions used in the study. Alternatively, the most robust approaches for generating systems with uniform behavior are those based on optimizing the SS criterion; these methods select systems without knowledge of experimental conditions and are therefore expected to yield devices which perform favorably across a reasonable range of experimental conditions. Consequently, both generation methods have their relative strengths. The Zhang *et al.* method allows one to generate systems with relatively uniform reaction rates, but its predictions depend strongly on both system design and experimental conditions. The SS criterion results in systems with less uniform reaction rates, but its predictions are robust to variations in both system design and experimental conditions. Ideally, future design methods will enable one to generate systems whose performance are both more uniform than the Zhang *et al.* method, and whose performances remain uniform under varying experimental conditions similarly to the SS criterion.

Section 1.9 – Improving the State-of-the-Art

So far, scientists prototyping new DNA systems have had relative success using the design process presented in Figure 1-4d above. However, virtually all have noted unexpected behaviors such as inconsistent formation yields, defective structures, poor performance, and/or total failure. It is plausible to hypothesize that at least some of these issues arise from the presence of non-target B-DNA structures, however there is both a

lack of tools for characterizing such structures and a lack of data demonstrating the impacts of such structures on system behavior.

In this dissertation, a new method for implementing DNA systems is presented (Figure 1-7 below). Briefly described, this process starts after the "abstract design" stage. This abstract design is translated into a domain-level design.[15] At this stage, each strand in the system is described using only binding domains and their complements. Each of the domains in this design are declared as either variable or fixed and given initial sequences. Next, the sequence of variable domains are manipulated in order to optimize a fitness metric. This process has been automated using the custom-written Sequence Evolver software (abbreviated SeqEvo). The quality of the system produced by SeqEvo is then scrutinized, and if necessary SeqEvo parameters are tuned and sequence optimization repeated. To automate this step, and to enable the evaluation of other existing systems, the Device Profiler (DevPro) software was created. If the quality of the system is deemed appropriate, then the process finishes and experimental characterization can be conducted.

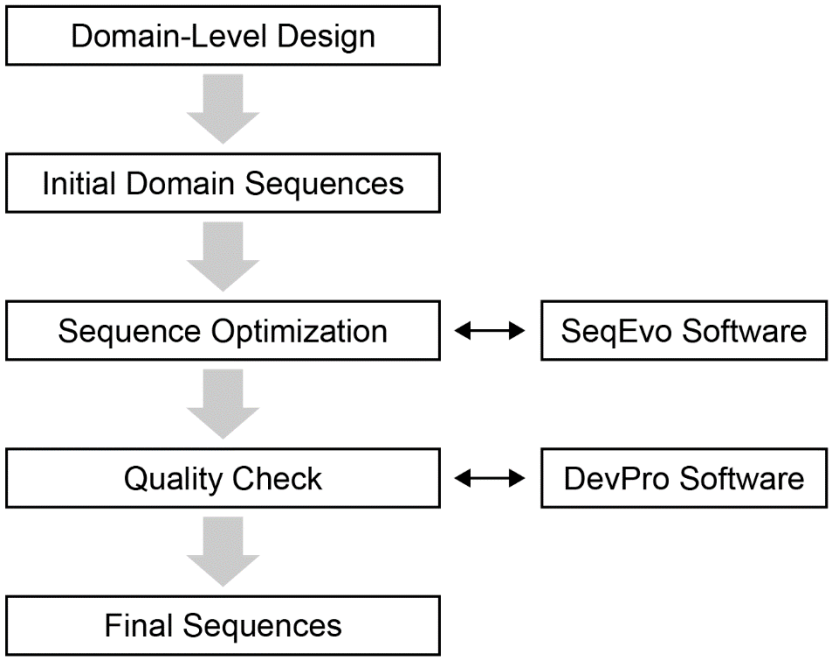


Figure 1-7 Visual summary of the proposed process for generating DNA systems. Two software tools have been created to help automate this process: the Sequence-Evolver (abbreviated SeqEvo) software for generating *in silico* optimized sequences, and the Device Profiler (abbreviated DevPro) software for generating detailed reports characterizing existing systems.

The work in the dissertation is composed of three major efforts, each of which is interdependent to the other two (Figure1-8 below). First, three new criteria for identifying systems with uniform behavior are proposed and studied (Figure 1-8 top). In order to evaluate the performance of these metrics, a robust heuristic algorithm for generating fit sequences was developed (Figure 1-8 right). Creation and optimization of this evolution-inspired algorithm became the second major effort. The development of the algorithm and fitness criteria necessitated the creation of two software tools (Figure 1-8 left). The creation of software tools efficient enough to both characterize and engineer large state-of-the-art DNA systems became the third major effort. Collectively, the new optimization criteria, the new sequence generation algorithm, and the efficient software implementation enabled the experimental characterization of systems with varying types

of non-target structures (Figure 1-8 center). This enabled us to test the hypothesis that small non-target B-DNA structures are responsible for previously observed kinetic variation. The results of this study both validate that the software/algorithm/criteria are functioning as intended and that their combination represents an improvement over current state-of-the-art methods.

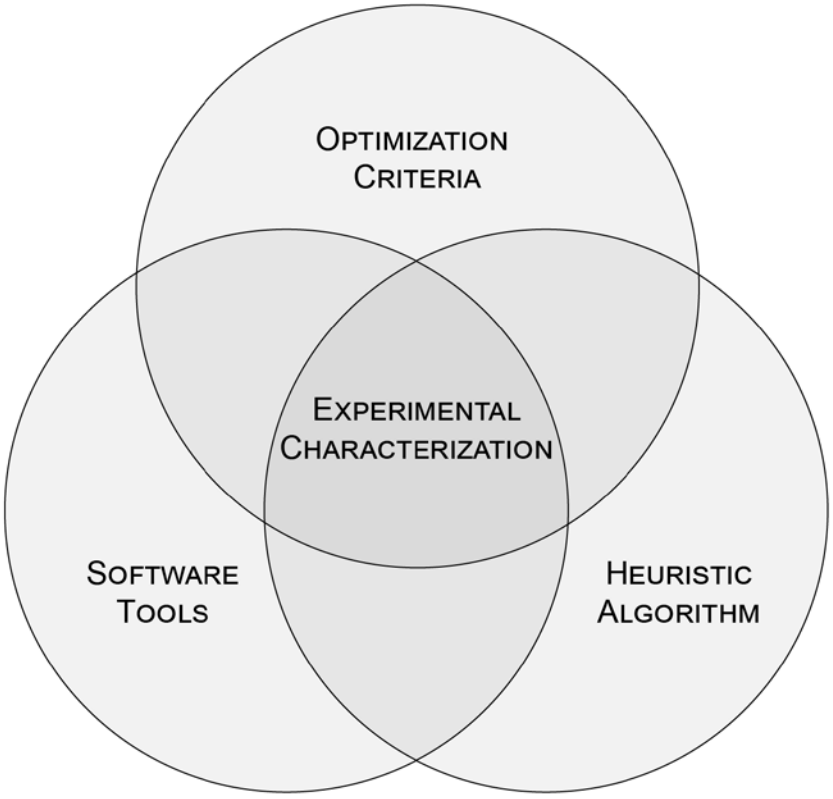


Figure 1-8 Key aspects of this dissertation and their interrelationships. (top) Three new optimization criteria were developed. (left) Two new software tools were created. (right) A new sequence generation algorithm was developed. (center) The optimization criteria, heuristic algorithm, and software tools were used to generate samples for experimental characterization.

CHAPTER TWO: QUANTIFYING SEQUENCE FITNESS

In order to reliably generate high-quality DNA systems via sequence optimization, it is first necessary to know what to optimize. Typically, one would like to have a property which is both simple for a computer to calculate and correlates strongly with a desired measure of performance. Such a property could serve as a metric or criterion for comparing the fitness of systems *in silico* and facilitate the automated generation of systems with a desired performance. Towards this goal, three new properties quantifying system fitness were proposed and investigated.

Section 2.1 – Quantifying Variations in Dynamic Behavior

One measure of dynamic behavior in DNA systems is the reaction rate of a specific target reaction. Consequently, one measure of behavior variation is the dispersion of these kinetic rates. Conveniently, recent scientific advancements such as the X-probe architecture[53] have enabled researchers to characterize the kinetics of many sequences implementing a single model system. This is demonstrated in two recent publications which studied the causes of kinetic variation and reported large sets of experimentally determined reaction rates.[36,37] In principle, this data can be used to evaluate the effectiveness of potential fitness criteria and refine the implementation process. Relative to the design process introduced in Figure 1-4d, this can be thought of as closing a feedback loop from experimental characterization to implementation (Figure 2-1a below).

From these two manuscripts, three experimentally-consistent datasets were extracted. In the first publication, Hata *et al.* demonstrated that thermodynamically unfavorable structures exhibit a marked impact on reaction rates.[37] To support this argument, the authors reported duplex formation rates for 47 implementations of a 23 bp DNA duplex (Figure 2-1b). This set of data was given the label "H25" based on the fact the data was collected at 25°C. The 47 rates reported were all measured under consistent experimental conditions and varied from 1.03×10^4 $M^{-1}s^{-1}$ to 4.8×10^6 $M^{-1}s^{-1}$ (Figure 2-1d). In the second publication, Zhang *et al.* created a model for predicting reaction rates based on similarity to the rates of already measured strand sequence.[36] The researchers reported duple-formation rates for 99 implementations of a 36 bp DNA duplex (Figure 2-1c). From this publication, two sets of data were extracted: 99 rate/sequence pairs recorded at 37°C (labeled "Z37"), and 96 rate/sequence pairs recorded at 55°C (labeled "Z55"). Rates in these data sets ranged from 1.6×10^4 to 2.5×10^7 and from 1.4×10^5 to 2.6×10^7 M^{-1} s^{-1} for the 37°C and 55°C data, respectively (Figure 2-1d). Between the two publications, methods applied for sequence generation and experimental characterization varied substantially. It is of specific note that the kinetic models and hence kinetic rates reported in the two studies differ.

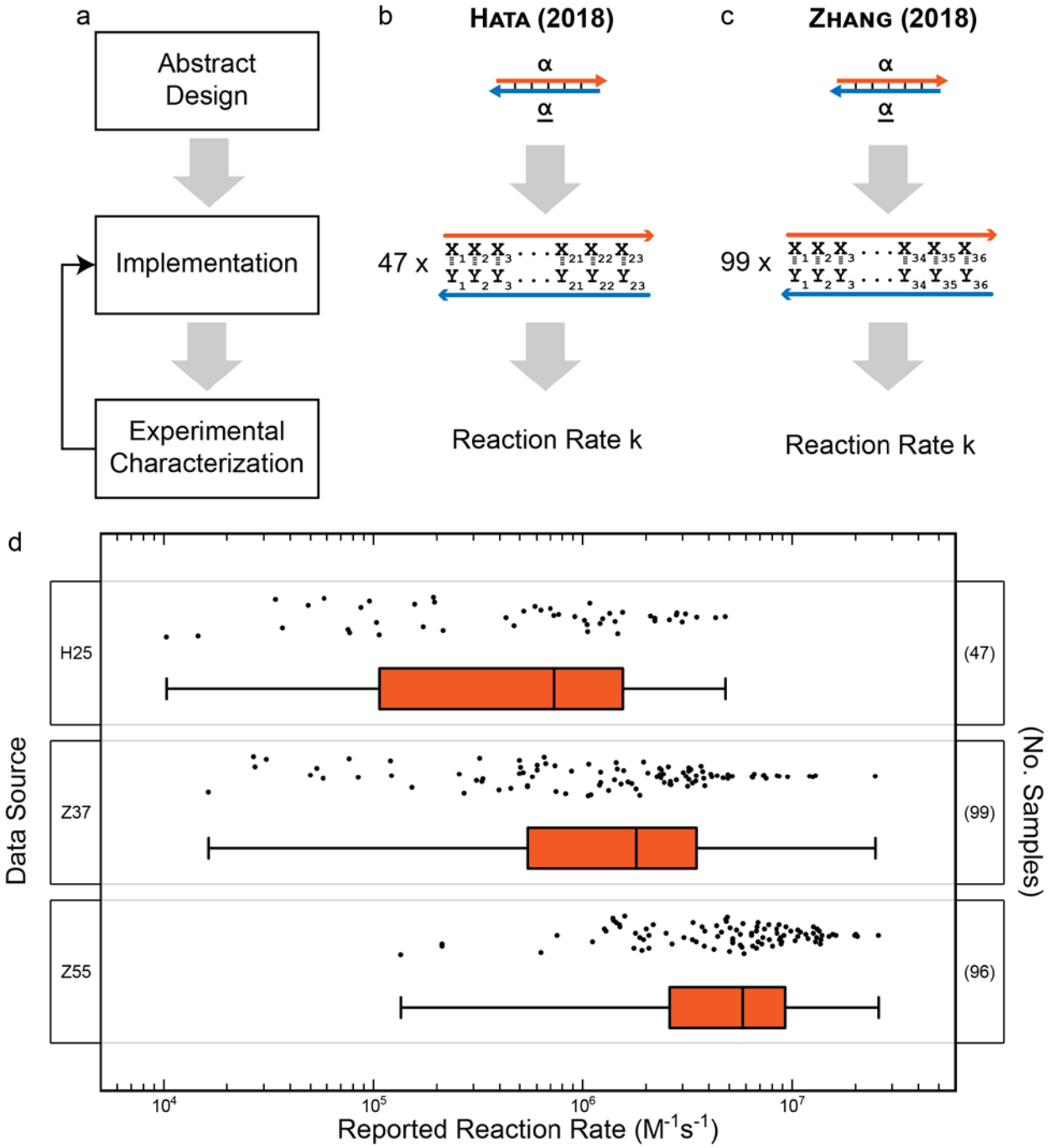


Figure 2-1 (a) Flow chart of typical design process with feedback from experimental characterization to implementation process highlighted. (b) Hata *et al.* studied the formation rates of 47 implementations of a 23-bp duplex at a temperature of 25°C.[37] (c) Zhang *et al.* studied the formation rates of 99 implementations of a 36 base-pair duplex at temperatures of 37°C and 55°C.[36] (d) Reaction rates reported in the three data sets presented as points and summarized by a median line, a box connecting the 25th and 75th percentiles, and bars connecting the min and max values.

The behavior of these simple model systems was observed to be highly variable and should be expected to create a host of issues for researchers designing structures from DNA. For structural systems such as DNA-Bricks or DNA Origami, this should lead to variation in strand incorporation, ultimately impacting both production rates and yields. For dynamic systems such as chemical reaction networks, this directly impacts device performance and presents as massive variation in reproducibility from implementation to implementation.

Section 2.2 – Quantifying Non-Target B-DNA Structures Using Network Fitness Score, Strand Fitness Score and Total Fitness Score

To quantify the presence of non-target B-DNA type structures, three fitness scores are proposed. Consider again the model system presented earlier (Figure 2-2a below). For this model system, the "complete" list of intermolecular and intramolecular B-DNA structures can be calculated by exhaustively considering all potential base pairings. These lists are "complete" in the sense that they contain all structures, including those that exist as a part of a larger structure. These lists can be summarized by binning structures based on their length and reporting the count of each structure-length (Figure 2-2b and c). The total number and length of structures are further quantified by assigning a score of 10^L points to each structure, where L is the structure length. This point assignment function ensures that a structure of length L receives the same score as ten structures with length $L\text{-}1$ for all lengths $L > 2$. System fitness is therefore biased toward having fewer, shorter structures. The sum of all intermolecular scores is termed the Network Fitness Score (NFS, Figure 2-2d) and the sum of all intramolecular scores is termed the Strand Fitness Score (SFS, Figure 2-2e). A weighted linear combination of the NFS and SFS is defined

as the Total Fitness Scores (TFS, Figure 2-2f). Based on these definitions, NFS can be thought of as a single number quantifying the ensemble of all non-target structures. For this number, lower values correspond to “better” systems and an NFS of zero corresponds to a hypothetical system containing no non-target structures. Similarly, SFS can be thought of as a number summarizing the ensemble of intramolecular non-target structures. Finally, since intramolecular structures are a subset of intermolecular structures, the TFS can be interpreted as a single number quantifying all non-target structures but with emphasis placed on the intramolecular structures. The intensity of the emphasis is controlled by the ratio of the two scoring weights (C_1 and C_2 in Figure 2-2f). Similar to both NFS and SFS, smaller TFS scores correspond with higher quality systems and a TFS of zero describes the ideal system containing no non-target structures.

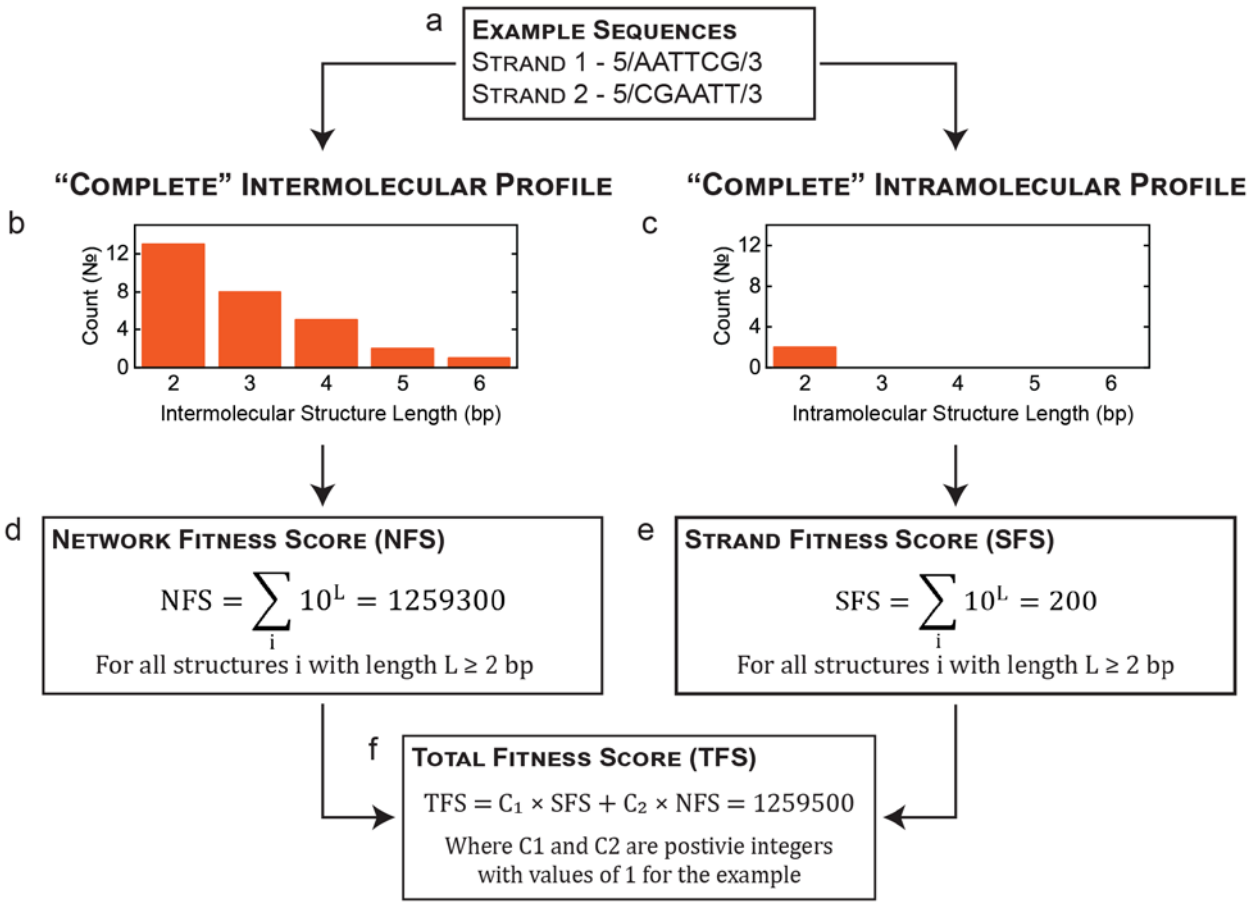


Figure 2-2 Calculation of Network Fitness Score (NFS), Strand Fitness Score (SFS), and Total Fitness Score (TFS). (a) Sequences from the model system presented in Figures 1-5. (b,c) Structural profiles summarizing the "complete" sets of intermolecular and intramolecular structures. (d,e) Calculation of NFS, SFS. (f) TFS is calculated as a weighted linear combination of SFS and NFS. By manipulating the C_1 and C_2 weights, TFS can be tuned to emphasize either NFS or SFS.

Section 2.3 – Do B-DNA Structures Explain Kinetic Variation?

The effectiveness of the proposed fitness scores were evaluated by using each to identify the "fittest" systems in the three published datasets. The distribution of kinetic rates for these fit systems were then compared to the remainder of the dataset. Kinetic variation and fitness score performance were quantified using the ratio of the Median to the Median Absolute Deviation (M/MAD). This metric quantifies the relative dispersion

in reaction rates in a manner which is both robust to outliers and aligned with the objective of engineering systems with uniform performance. For this metric, larger values correspond with better kinetic reproducibility.

Method and Results

The gathered sequence/rate pairs were used to evaluate the effectiveness of four fitness scores (Figure 2-3a below). For this purpose, the DevPro software was created. DevPro accepts a domain-level design and associated domain sequences as input, assembles the strand sequences for the system, calculates the set of all non-target structures, and then finally calculates the SFS, NFS, and TFS for the system. The software was used to analyze the sequences in each dataset and rank them in order of decreasing quality. From the list of ranked sequences, the five fittest sequences (and any sequences of similar quality) were identified (Figure 2-3b). The kinetic rates reported for these fittest sequences were statistically analyzed using P-values and M/MAD ratios.

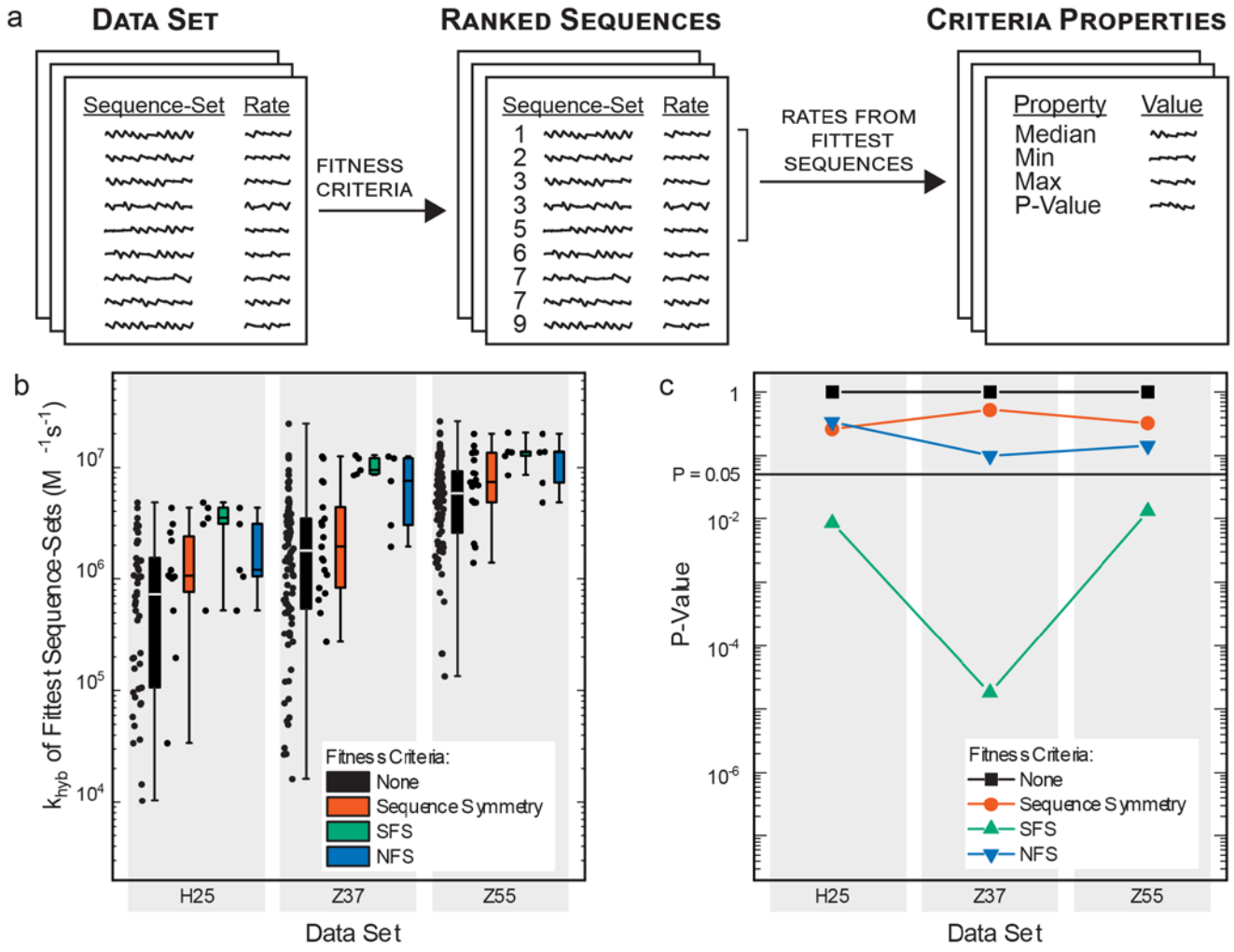


Figure 2-3 Effectiveness of different fitness criteria at identifying kinetically uniform sequences. (a) Cartoon of the criterion evaluation method. (b) Hybridization rates of the fittest systems within each dataset, as judged by one of the four criteria. Three datasets were analyzed: measurements by Hata *et al.*[37] at 25°C (H25), measurements by Zhang *et al.*[36] at 37°C (Z37), and measurements by Zhang *et al.*[36] at 55°C (Z55). (c) P-values calculated by comparing the fittest systems to the associated general population (labeled "none").

To test the null hypothesis that the rates of the fittest sequences were drawn from the same distribution as the general population, a two-sample Kolmogorov-Smirnov test was used. This test provides a non-parametric method for comparing two populations to determine the likelihood they were drawn from the same sample. P-values resulting from this test represent the likelihood that the two sets of samples were drawn from the same

distribution. In this study, the resulting P-values (Figure 2-3c above) indicate that only sequences which are SFS-fit are reliably distinct from the general population. Interestingly, neither the state-of-the-art Sequence-Symmetry nor the NFS selected sequences produced kinetically distinct populations. This was interpreted as evidence that the majority of kinetic deviations in these systems arise from intramolecular B-DNA type structures. While it is likely that the structures leading to kinetic deviation are intramolecular in nature, each intramolecular structure logically implies the existence of intermolecular structures. As such, it is important to note that this is evidence of correlation but not necessary causation. Consequently, it is clear that systems containing fewer intramolecular structures have more uniform kinetics, but it is not necessarily clear why.

These findings are reinforced by the M/MAD ratios calculated for each set of fit systems and reported in Table 2-1 below. The reaction rates of SFS-fit sequences were observed to be both larger and more uniform than the remaining sequences. Prior to filtering, the datasets were observed to have M/MAD ratios ranging from 1.15 to 1.68. In comparison, SFS-fit systems were observed to have M/MAD ratios ranging from 4.38 to 13.3. The SS-fit and NFS-fit systems were observed to reliably result in ratios larger than the unfiltered population, but smaller than SFS. These M/MAD ratios are a measure of the relative dispersion such that larger values are more favorable and correspond to a narrower rate distribution. Due to its definition, the inverse of the M/MAD ratio can be thought of as a fraction or percentage of the median such that most rates exist within plus or minus this fraction of the median rate. For example, the SFS-fit systems in the H25

dataset were observed to have an M/MAD ratio of 4.38, corresponding to typical kinetic variation of ± 22.8%.

Table 2-1 Select properties of the fittest systems in each criterion/dataset combination. Reported values include: the duplex-formation rate kDF, the median rate (M), the Median-Absolute-Deviation of rates (MAD), and P-values calculated using a two-sample Kolmogorov-Smirnov test.

Data Set	Fitness Criterion	k_{DF} ($M^{-1}s^{-1}$) M ± MAD	M/MAD	P-Value
H25 [37]	None	$(7.30 \pm 6.3) \times 10^5$	1.15	1.00
	SS	$(1.07 \pm 0.71) \times 10^6$	1.51	0.264
	SFS	$(3.50 \pm 0.80) \times 10^6$	4.38	0.00842
	NFS	$(1.21 \pm 0.69) \times 10^6$	1.76	0.343
Z37 [36]	None	$(1.76 \pm 1.42) \times 10^6$	1.24	1.00
	SS	$(1.96 \pm 1.26) \times 10^6$	1.56	0.530
	SFS	$(9.53 \pm 1.05) \times 10^6$	9.07	1.81×10^{-5}
	NFS	$(7.51 \pm 4.63) \times 10^6$	1.62	0.0990
Z55 [36]	None	$(5.74 \pm 3.42) \times 10^6$	1.68	1.00
	SS	$(7.34 \pm 2.51) \times 10^6$	2.92	0.323
	SFS	$(1.37 \pm 0.10) \times 10^7$	13.3	0.0131
	NFS	$(1.36 \pm 0.63) \times 10^7$	2.18	0.143

In addition, the effectiveness of TFS was determined for several combinations of weighting parameters (Figure 2-4). TFS's with C_1/C_2 ratios from 10^{-6} to 10^6 were studied. It was observed that the TFS of the fittest sequences were approximately equal to NFS when ratios were smaller than 0.1 and approximately equivalent to SFS when ratios were larger than 10,000. TFS was observed to be effective on all datasets for C_1/C_2 ratios

greater than 100, and to be most effective for ratios greater than 10,000. This result can again be explained by a strong correlation between intramolecular B-DNA type structures and kinetic variation.

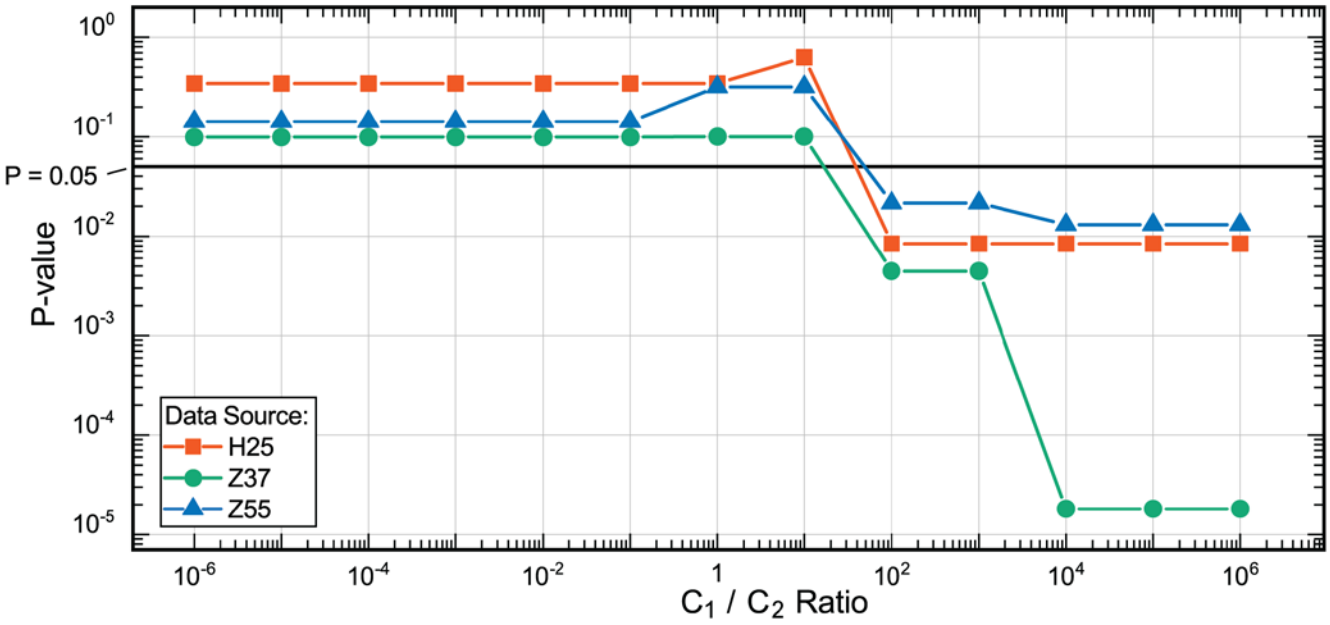


Figure 2-4 Tuning of the Total Fitness Score (TFS) weighting parameters to achieve statistically significant P-values. Three datasets were analyzed: measurements by Hata *et al.*[37] at 25°C (H25, orange squares), measurements by Zhang *et al.*[36] at 37°C (Z37, green circles), and measurements by Zhang *et al.*[36] at 55°C (Z55, blue triangles). P-values are the result of a two-sample Kolmogorov-Smirnov test comparing the fittest systems selected by the TFS function with the associated general population.

Discussion

The SFS and TFS with certain weighting parameters are observed to reliably identify statistically significant subpopulations across the three data sets (P-value < 0.05). These populations have rates which are larger and less disperse than the remainder of the populations, suggesting that intramolecular non-target B-DNA structures are a key contributor to kinetic variation. Kinetic reproducibility was quantified using the M/MAD ratio, and the larger ratios observed for each of these populations indicates improved

kinetic reproducibility for SFS-fit and certain TFS-fit systems. The M/MAD values for the SFS systems range between 4.4 and 13, corresponding with kinetic variations of ± 23% and ± 7.7%, respectively. Interestingly, neither the state-of-the-art SS criterion nor NFS selected systems were statistically distinct from the unfiltered general population.

Section 2.4 – Are TFS-Fit Sequences Kinetically Uniform?

While a strong correlation between intramolecular structures and kinetic variation was observed, the analysis of existing data does not necessarily imply that new systems created using these principles will be kinetically uniform. To confirm that engineering systems with minimal intramolecular and limited intermolecular B-DNA type structures leads to kinetic uniformity, twelve new systems were generated and experimentally characterized.

Methods and Results

System Generation

For the purpose of generating fit systems, the SeqEvo software was created. SeqEvo utilizes an evolution-inspired heuristic algorithm to identify systems which are TFS-fit. The TFS weighting parameters supplied to the program can be manipulated to emphasize either SFS-fit or NFS-fit systems. As a design for the model system, three strands capable of undergoing two target reactions were identified (Figure 2-5a below). Two of these strands are fully complementary and contain 41 variable bases (represented with X's and Y's such that X_i pairs with Y_i in Figure 2-5). This model system design is intended to form two target structures (Figure 2-5b): a fully complementary B-DNA duplex composed of Strands 1 and 2 (referred to as complex 1:2) and a partially complementary B-DNA duplex composed of Strands 2 and 3 (referred to as complex

2:3). The design is intended to undergo two target reactions. In the first target reaction, Strand 1 and Strand 2 react to form a 49 base-pair B-DNA structure (Figure 2-5c). In the second reaction, Strand 1 interacts with the single-stranded region in complex 2:3 to displace strand 2 and form the 1:3 complex (Figure 2-5d). This mechanism is known as toehold-mediated strand displacement.[38] The design is such that the 8 bases utilized in this reaction are of fixed sequence (5-TCTCCATG-3 and 5-CATGGAGA-3). This was done to eliminate kinetic variation in the strand-displacement reaction known to occur based on toehold sequence.[27]

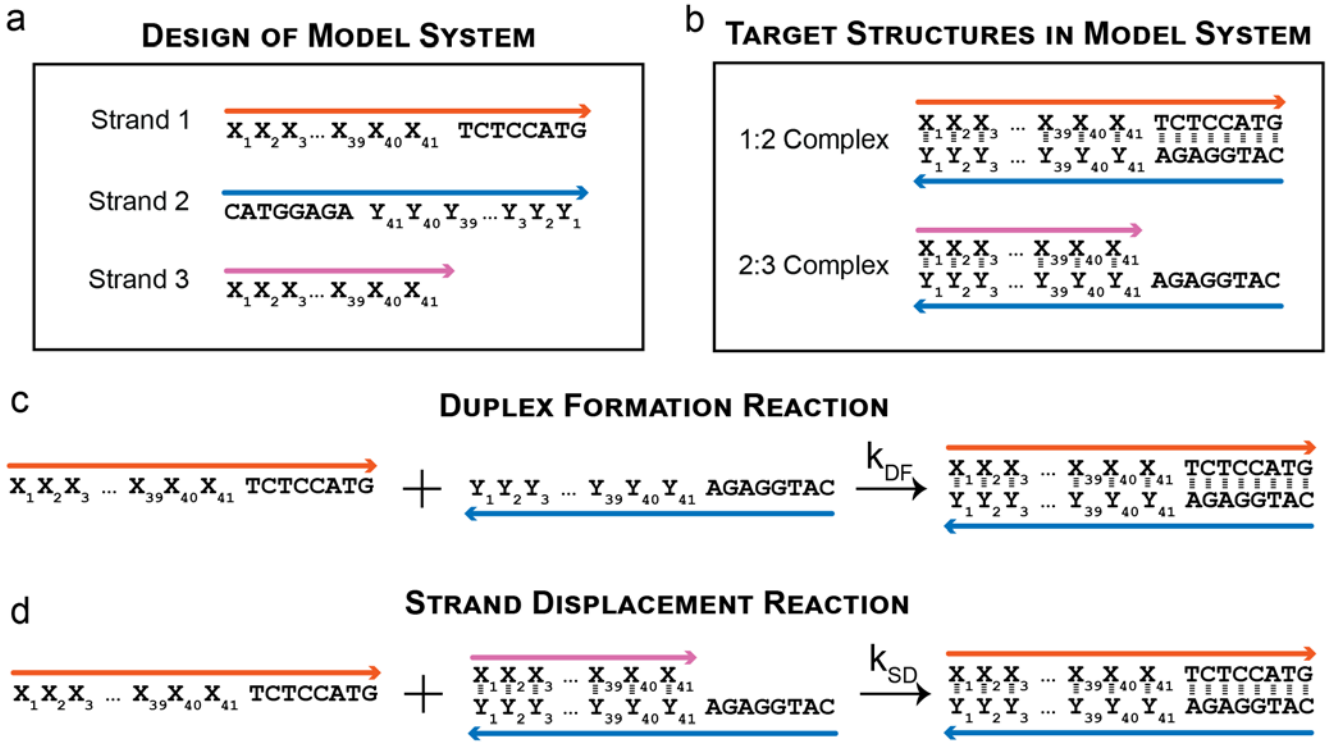


Figure 2-5 A model system for studying the impact of non-target structures on reaction kinetics. (a) Three strands compose the model system. (b) The two target structures in the model system. (c) Schematic of the duplex-formation target reaction. (d) Schematic of the strand-displacement target reaction.

Twelve systems implementing this design were generated. The variable sequences (X's in Figure 2-5 above) for each generated system are reported in Figure 2-6 below.

From these variable sequences, strands were compiled and are reported in Appendix A.1. The generated systems were organized into four design groups: three implementations generated via a pseudo-random number generator (*RND* group), three implementations with the TFS property optimized (*TFS* group with $C_1 = 1$, $C_2 = 1$), three implementations with the SFS property optimized (*SFS* group with $C_1 = 1$, $C_2 = 0$), three implementations with the NFS property optimized (*NFS* group with $C_1 = 0$, $C_2 = 1$). These weighting parameters were chosen based on a binary on/off mentality intended to identify whether (a) these TFS weights would be effective for the target system and (b) whether optimization of solely SFS or solely NFS would result in kinetic reproducibility. It is important to note that the SeqEvo software results in sequences with two relevant properties. First, all 12 variable sequences are composed of the same bases arranged into different sequences, meaning that A/T/C/G content in each domain is identical. Second, The SeqEvo software does not allow for sequences containing stretches of A's C's T's or G's longer than a threshold set by the user. For these implementations the thresholds were set at 6, 3, 6 and 3 bases, respectively, based on previously successful design methods.[15,29]

Design Group	System Name	Bases X_1 … X_{41}
RND	RND-1	GTGTCAACACCTCGCTAGAGATGGTGCGCTAAATTACGCT
	RND-2	GATTAGTCATTAAGGGATCGACACCACGGGCTTCTTCCGA
	RND-3	TCCTATGTACAGTCGTACGGACTATTGCGGAACCCTGAGA
TFS	TFS-1	CCAATCGCCCGTCGTAGGTGTGTCAGTAATAAAGCAGTTC
	TFS-2	AGTGTATCCAAAGCCCGTAAGTCGCAGGTTCGTGTCAATC
	TFS-3	CGTAGTGTGTCAGCAAAGTCCAATAGGTTCGCCCGTAATC
SFS	SFS-1	AAAAGTGTGTAAAAAAGTCCCGTGTCCGTGTGTCCGTCCC
	SFS-2	CGTGTGTGTGTCCCGTAAAAGTAAAAAAGTCCCGTGTCCC
	SFS-3	GTGTAAAAGTGTCCCGTGTCGTAAAAAAGTCCCGTGTCCC
NFS	NFS-1	TATCGTCACAGTTCGGTTCCAAAGGGCAATCAGCGTAGTC
	NFS-2	CGGCGTAAGCAATAGGTTTCACAATCCCAGGTAGTCGTTC
	NFS-3	GTAAATCCCGTGCTAAAGTATCGTCGCCAAGGTTCAGGTC

Figure 2-6 Variable sequences for the twelve systems generated to implement the model system design presented in Figure 2-5. A full list of strand sequences is provided in appendix A.1.

Generated Systems

The non-target structures in twelve systems were characterized using DevPro and are reported as interference profiles in Figure 2-7 below. The resulting structural profiles were found to be categorizable into three distinct shapes: (1) implementations generated in the NFS and TFS design groups, which contain neither intramolecular structures longer than 2 bp or intermolecular structures longer than 4 bp; (2) implementations generated in the SFS design group, which contain no intramolecular structures longer than 2 bp, but intermolecular structures up to 10 bp in length; and (3) implementations generated in the RND design group, which contain substantial numbers of both intramolecular and intermolecular structures (up to 5 bp and 8 bp, respectively). It was noted that while the shape of the TFS and NFS profiles are similar, sequences in the TFS

design group contain, on average, approximately 10% fewer intramolecular structures than the NFS design group.

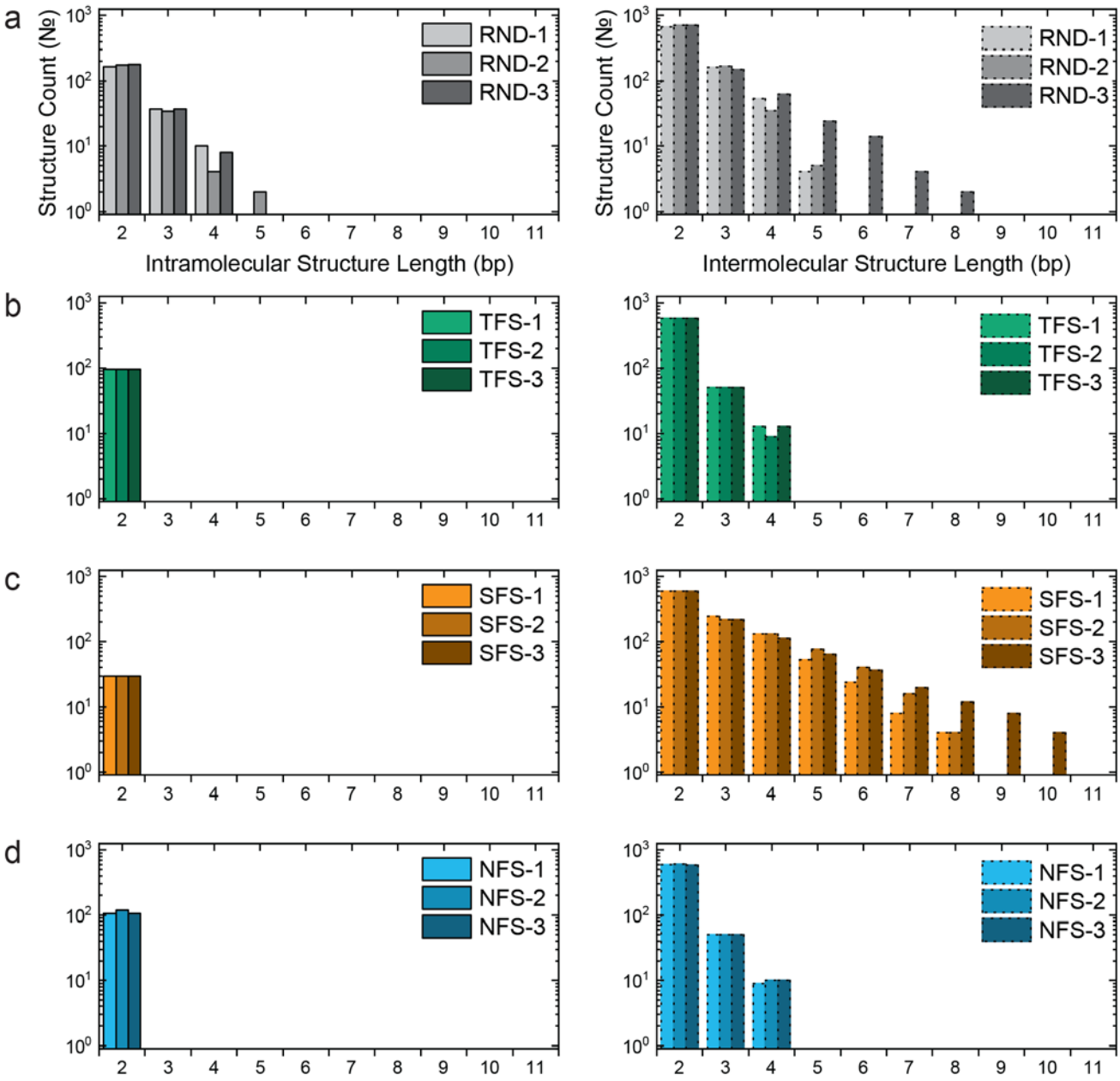


Figure 2-7 Structural profiles of the twelve generated systems. (a) Intermolecular (left) and intramolecular (right) profiles of the three randomly generated systems. (b) Profiles of the three TFS-fit sequences. (c) Profiles of the three SFS-fit sequences. (d) Profiles of the three NFS-fit sequences. These structural profiles are complete in the sense that they contain all non-target structures, including those which exist within a larger structure.

Experimental Characterization

The behavior of the twelve generated systems were characterized by monitoring reactant concentration in real time using fluorescence measurements. This technique is illustrated in Figure 2-8 below using the TFS-1 system (Figure 2-8a) at 20°C as an example. First, strands were purchased from Integrated DNA Technologies (www.idtdna.com) with dye/quencher modifications and high-performance liquid chromatography purification. The three modifications were: a 5' Cy3 dye for strand-1, a 5' Cy5 dye for Strand-3, and a 3' "Black Hole Quencher 2" (BHQ2) for Strand-2. Reactants were prepared in 1x TE buffer supplemented with 1 M NaCl (Figure 2-8b). Reactants were prepared such that samples were 3 ml of 10 nM reactants in a 1 cm disposable cuvette. Cuvettes were pre-treated with "Superblock" blocking buffer to prevent DNA adhesion. Sample fluorescence was monitored using one of two Cary Eclipse spectrophotometers (Figure 2-8c). Samples were excited at 548 nm and emission was monitored at 573 nm based on the excitation/emission spectra of the Cy3 dye. Reactants were allowed to come to thermal equilibrium with a temperature-controlled sample holder prior to experiments, and the temperature was then held constant during the experiment via the sample holder. Each sample began with only the Cy3 labeled Strand-1 present at a concentration slightly greater than 10nM. After thermal equilibrium was established, the sample was removed from the holder. The second reactant was injected, and the sample was mixed using a pipette. The sample was then returned to the sample holder and fluorescence monitored. Sample cooling during the mixing process was observed to be negligible. Detailed reports of the fluorescent measurements for each system/temperature/reaction combination are provided in Appendix A.2.

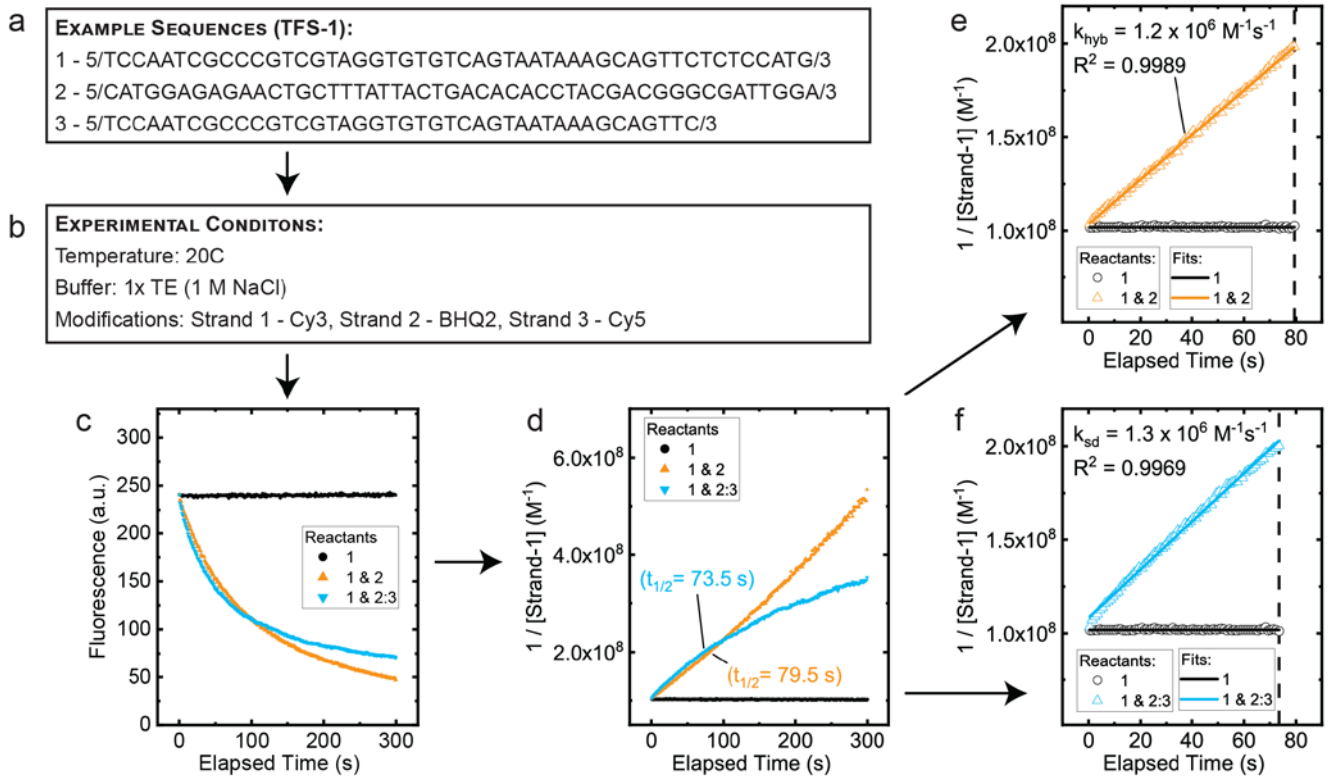


Figure 2-8 Method for experimentally characterizing reaction kinetics. (a) One of the generated systems for use as an example. (b) Dye/quencher functionalized strands were prepared at pre-determined experimental conditions. (c) Sample fluorescence was monitored in real-time. (d) Plot of inverse strand concentration vs elapsed time. (e,f) Linear fits applied to all data preceding reaction half completion in d.

Kinetic Modeling

Both the duplex-formation and strand displacement reactions (Figure 2-5c,d) were modeled as bimolecular and irreversible (equation 1 below).

$$A + B \xrightarrow{k} C \qquad 1$$

For this reaction, the law of mass action dictates that the rate of reactant consumption and the rate of product creation are equal (eq. 2).

$$\frac{d[A]}{dt} = \frac{d[B]}{dt} = -\frac{d[C]}{dt} = -k[A][B] \qquad 2$$

As such, the assumption that reactants are initially at equal concentrations (stoichiometry) implies that the reactant concentrations remain equal indefinitely.

$$[A]_0 = [B]_0 \quad (3)$$

$$[A] = [B] \quad (4)$$

Substitution and rearrangement of eq. 2 allows separation of variables in eq. 5.

$$-\frac{d[A]}{[A]^2} = k\,dt \quad (5)$$

Integration of eq. 5 yields a linear relationship between the inverse reactant concentration and time.

$$\frac{1}{[A]} = kt + C \quad (6)$$

Based on this model, one can expect a plot of inverse concentration vs time to be linear for any stoichiometric, bimolecular, and irreversible reaction. For both duplex formation and strand displacement reaction, these plots were observed to be approximately linear for times prior to half completion (Figure 2-8d above). Nonlinear behavior was observed beyond half completion. This is consistent with the increasing deviation from stoichiometry which is expected for such reactions. The slope of the linear region is equal to the bimolecular rate constant k, which was extracted using a linear fit to the data (Figure 2-8e, f).

Measured Reaction Rates

One hundred and fifty-two total reaction rates were experimentally determined. For each of the twelve implementations (RND-1, RND-2, RND-3, TFS-1, TFS-2, TFS-3, SFS-1, SFS-2, SFS-3, NFS-1, NFS-2, and NFS-3), the rate of both target reactions (k_{DF} and k_{SD}) were measured at six temperatures (10°C, 20°C, 30°C, 40°C, 50°C and 60°C).

Sample preparation and rate measurements were repeated two additional times for the TFS-3 system at 20°C and 40°C for both k_{DF} and k_{SD}. The resulting k_{DF} measurements of 6.1 x 10^5, 7.0 x 10^5, 6.7 x 10^5 $M^{-1}s^{-1}$ at 20°C and 5.9 x 10^6, 6.0 x 10^6, and 5.8 x 10^6 $M^{-1}s^{-1}$ at 40°C indicate the precision of this method is such that a single measurement of each rate is reasonably appropriate for the target study. The resulting k_{SD} measurements of 1.5 x 10^6, 1.8 x 10^6, 1.5 x 10^6 $M^{-1}s^{-1}$ at 20°C and 1.6 x 10^6, 1.6 x 10^6, and 1.7 x 10^6 $M^{-1}s^{-1}$ at 40°C indicate similar precision for the strand-displacement rates.

The 12 sets of rates determined for the duplex-formation reaction are reported in Figure 2-9. These rates were observed to span five orders of magnitude, with rates ranging from a minimum value of 9.6 x 10^3 $M^{-1}s^{-1}$ (RND-1 at 10 °C) to a maximum value of 8.0 x 10^7 $M^{-1}s^{-1}$ (TFS-1 at 60 °C). The largest range observed at a given temperature resulted from the measurements at 10 °C, which spanned four orders of magnitude from 9.6 x 10^3 $M^{-1}s^{-1}$ to 3.7 x 10^6 $M^{-1}s^{-1}$ (RND-1 and SFS-1, respectively). Duplex formation M/MAD ratios were calculated for each design-group at each temperature yielding average ratios of 1.5, 19.7, 5.5, and 4.4 for the RND, TFS, NFS and SFS groups, respectively. The largest duplex formation M/MAD ratio observed was a value of 44 (TFS group at 30 °C). The smallest M/MAD ratio observed was a value of 1.4 (RND group at 40 °C). The majority of rate/temperature profiles were observed to be well described by an Arrhenius model (equation 7 and Figure 2-10 below) relating the kinetic rate (k) to an activation energy (E_a), a pre-exponential factor (A), the Boltzmann constant (k_b), and the temperature (T).

$$k = A \exp\left(\frac{-E_a}{k_b T}\right) \qquad 7$$

Based on this observation, application of a linear fit to the plot of the natural logarithm of

reaction rates as a function of inverse temperature yielded the Arrhenius slope, intercept and R^2 values reported in Table 2-2 below.

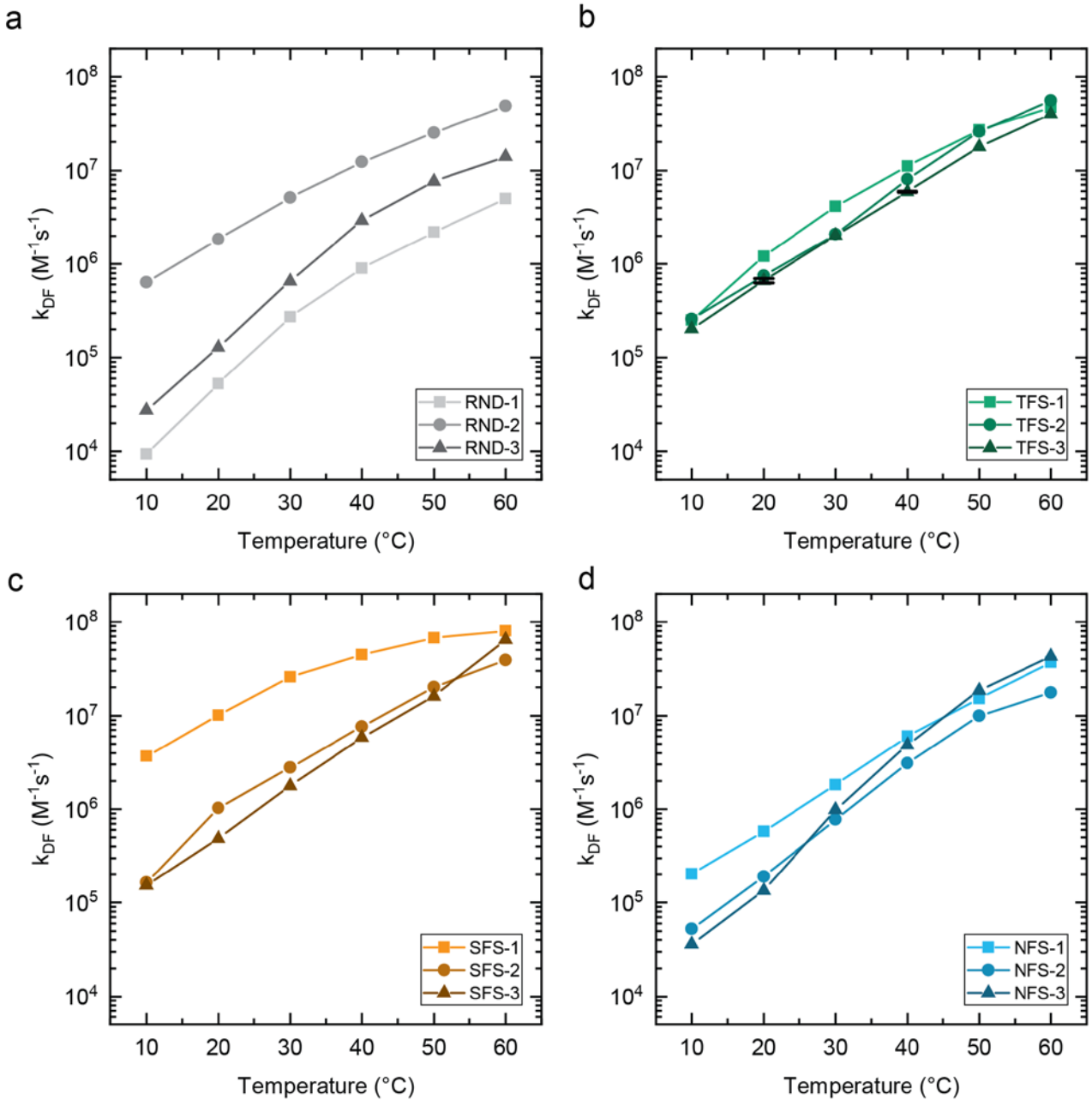


Figure 2-9 Experimentally determined duplex-formation (k_{DF}) rates for the twelve implementations of the model system. The discrete data points are connected by lines to aid the eye. Experiments were performed in triplicate for the TFS-3 system at 20 °C and 40 °C. The error bars on these data points span from the mean to the standard deviation of the three measurements.

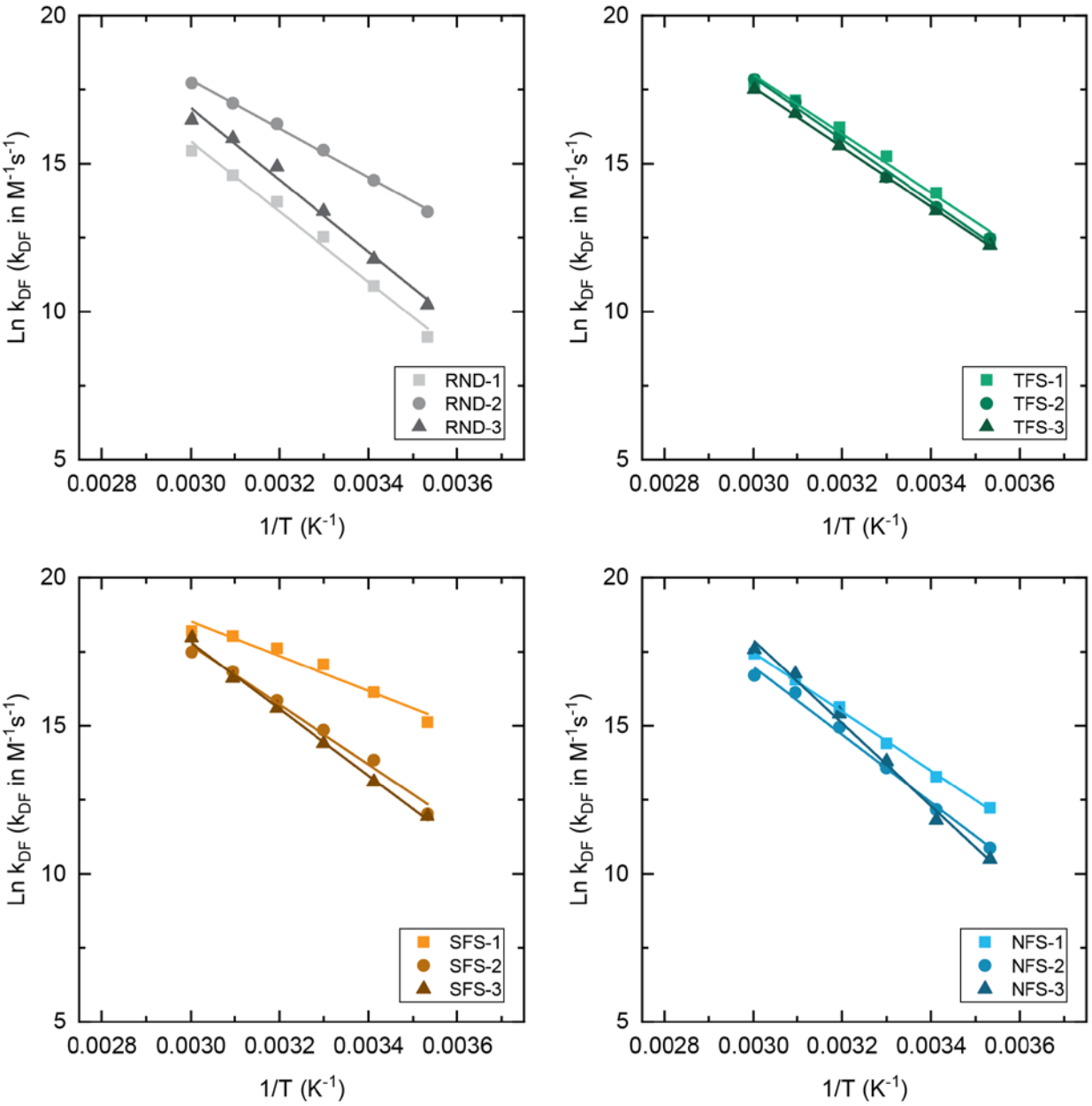


Figure 2-10 Arrhenius fits to the experimentally determined duplex-formation (k_{DF}) rates for the twelve implementations of the model system. Discrete data points are shown as symbols, with lines illustrating a linear fit to the data.

Table 2-2 Arrhenius parameters extracted from the duplex-formation rates (kDF) of each implementation. The activation energy (Ea), pre-exponential factor (A), and R2 of the Arrhenius fit are reported.

	Arrhenius Parameters		
Sequence Set	E_a (10^{-19} J)	A ($M^{-1}s^{-1}$)	R^2
RND-1	1.71	1.41×10^{23}	0.995
RND-2	1.10	1.42×10^{18}	0.985
RND-3	1.68	1.72×10^{23}	0.989
TFS-1	1.38	7.75×10^{20}	0.993
TFS-2	1.36	3.15×10^{20}	0.992
TFS-3	1.42	1.16×10^{21}	0.998
SFS-1	0.654	1.18×10^{14}	0.861
SFS-2	1.24	1.73×10^{19}	0.962
SFS-3	1.50	7.73×10^{21}	0.997
NFS-1	1.28	4.27×10^{19}	0.993
NFS-2	1.57	1.56×10^{22}	0.992
NFS-3	1.82	6.94×10^{24}	0.988

The strand-displacement rates measured for each device are reported in Figure 2-11 below. Strand displacement reactions were observed to systematically deviate from the bimolecular model such that the reactant consumption slowed as elapsed time increased. This is evident in the Appendix A.2 graphs where the model is shown as a blue line and the strand-displacement reaction is shown as green squares. It is likely that a more complex model such as the three-step model proposed by Zhang and Winfree would better describe strand-displacement kinetics.[27] However, the advantages of the

simple bimolecular model are that it is based on a single kinetic rate and adequately quantifies system behavior such that behavior variation can be studied. More specifically, if systems have uniform dynamic behavior one would expect them to be have similar apparent bimolecular rates, regardless of the fact that model is an over simplification.

Measured strand displacement rates were observed to be less variable than the duplex-formation rates; the maximum and minimum rates spanned 3 orders of magnitude and ranged from 1.9×10^{4} $M^{-1}s^{-1}$ (RND-1 at 10 °C) to 1.9×10^{6} $M^{-1}s^{-1}$ (SFS-2 at 30 °C). This trend can potentially be explained by two factors: (1) several bases in the 2:3 complex are already in a B-DNA type structure, potentially eliminating their contribution to kinetic variation, and (2) the strand displacement reaction is designed to proceed through a specific reaction pathway including toehold formation, potentially eliminating kinetic variation arising from alternative reaction pathways. Strand displacement M/MAD ratios were calculated for each design-group at each temperature resulting in average ratios of 2.2, 9.1, 6.8, and 8.3 for the RND, TFS, NFS and SFS design groups, respectively. The largest M/MAD ratio observed was a value of 28 resulting from the TFS design group at 50 °C. The smallest M/MAD ratio observed was a value of 1.4 (RND group at 40 °C).

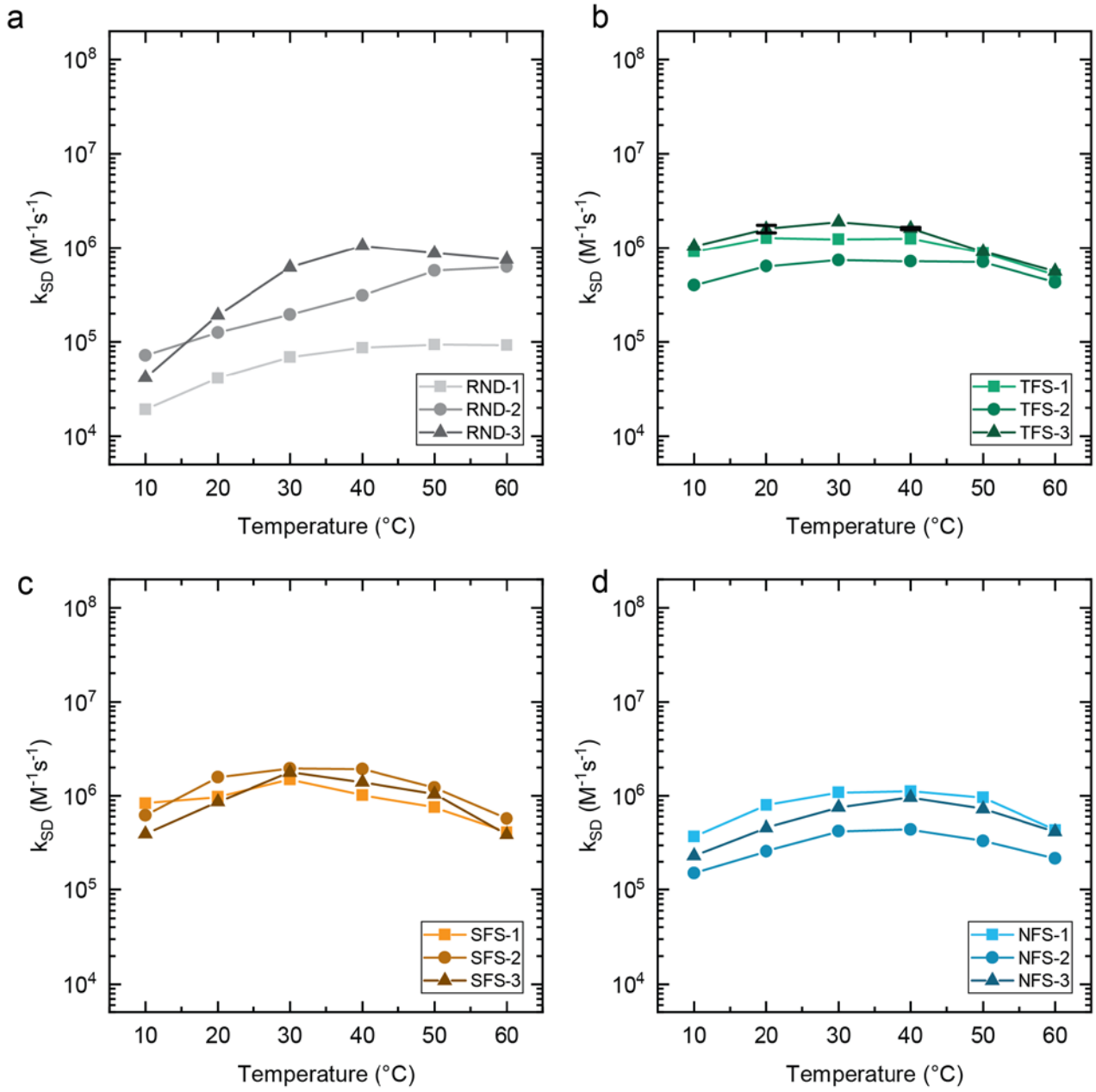


Figure 2-11 Experimentally determined strand-displacement rates for the twelve implementations of the model system. Experiments were performed in triplicate for the TFS-3 system at 20 °C and 40 °C. The error bars on these data points span from the mean to the standard deviation of the three measurements.

Discussion

Several trends in the experimentally characterized rates provide important insight into the relationship between non-target structures and kinetic variation. First, rates for both reactions are observed to be highly sequence-dependent. Indeed, it is observed that variation of sequence alone leads to rates spanning up to four or three orders of magnitude for the duplex-formation and strand-displacement reactions, respectively. For the duplex-formation reaction, kinetic variations of this magnitude have been observed previously, with the data reported by Hata *et al.* and Zhang *et al.* similarly spanning up to four orders of magnitude.[36,37] In addition, our observation of strand-displacement rate variation is consistent with the variations observed by Olson *et al.* while studying the impacts of sequence variation on chemical reaction network dynamics.[54] In addition, a study by Zhang and Winfree demonstrated that variation in toehold sequence and size can lead to strand-displacement rates varying up to seven orders of magnitude.[27] The results of this study expand upon this finding, making it clear that even systems with fixed toeholds vary by up to three orders of magnitude. This type of variation may also help explain the deviations from the three-step model observed by Zhang and Winfree for toeholds with high thermodynamic stability.

For the duplex formation reaction (Figures 2-5c and 2-9), rates were observed to be Arrhenius for all systems, and most reproducible for the TFS design group. The relative uniformity of the Arrhenius parameters for the TFS design group suggests that that both intra- and inter-molecular non-target structures exhibit an influence on reaction kinetics. This observation is based on the fact that neither the SFS nor the NFS design groups exhibited the same temperature profiles as the TFS group. In addition, the

observation that most systems exhibit similar Arrhenius behavior suggests that reaction mechanisms are largely preserved regardless of non-target structures. These findings are consistent with the nucleation-and-zipper model of duplex-formation described by equation 8 below.[55] In this model, nucleated intermediates form based on the bimolecular rate k_{1f}. These nucleated intermediates either dissociate back into reactants or proceed to reaction completion based on the unimolecular rate constants k_{1r} or k_2, respectively. Based on the steady-state approximation, such a reaction results in an apparent bimolecular kinetic rate (k_{app}) described by equation 9. Insufficient evidence is observed in the data to speculate if the Arrhenius barrier observed for bimolecular duplex-formation rates arises from a single or multiple Arrhenius barriers in k_{1f}, k_{1r} and k_2. Interestingly, both Arrhenius[56-59] and non-Arrhenius[37,56,60] temperature/rate profiles have been reported in the literature. However, non-Arrhenius behavior may be explainable based on the use of chemical buffers with relatively low ionic concentrations (only NaCl present and at concentrations < 0.2 M). Data presented by Wallace *et al.* appears to directly demonstrate a transition from non-Arrhenius to Arrhenius behavior based on increasing ionic concentration.[56] This can be potentially explained by two theories: (1) The reaction mechanism is impacted by the reactant charge and these effects decrease as the ionic concentration is increased. This suggests that reaction kinetics are Arrhenius in the absence of charge effects, and non-Arrhenius in their presence. (2) Alternatively, the reaction mechanism may be non-Arrhenius by nature and increasing ionic concentration could create Arrhenius behavior based on a cage-effect mechanism.[61]

$$A + B \underset{k_{1r}}{\overset{k_{1f}}{\rightleftarrows}} AB^{\dagger} \xrightarrow{k_2} AB \quad (8)$$

$$k_{app} = \frac{k_{1f}}{1 + \frac{k_{1r}}{k_2}} \quad (9)$$

The values of the Arrhenius activation energy (E_a) and pre-exponential factor (A) describing the duplex-formation rates of the 12 implementations are observed to be non-independent and strongly correlated (Figure 2-12 below). The correlation is such that a plot of the natural logarithm of A as a function of E_a appears linear in nature (eq. 10 below). Such relationships in DNA have been previously reported in the literature, and were interpreted as a consequence of the underlying linear free energy relationship.[57] This observation was confirmed using a linear fit (red line in Figure 2-12) resulting in an R^2 value of 0.9889, an intercept of 19.4 (a in eq. 10 below), and a slope of 2.05 x 10^{20} (b in eq. 11). Following the combination of equations 1 and 3, the declaration of constants C and T_c (eq. 11), and algebraic rearrangement, an empirical kinetic model can be derived (eq. 12). This model suggests that the duplex formation rates of the 12 devices are largely explainable based on two variables (E_a and T), and three constant parameters (C, T_c, and k_b). One interesting feature of this kinetic model is the critical temperature parameter T_c, which can be interpreted as a hypothetical critical temperature at which device kinetics should be uniform and equal to the pre-exponential constant C. The linear fit of the Arrhenius parameters predicts values of 82 °C and 2.7 x 10^8 $M^{-1}s^{-1}$ for T_c and C, respectively. Non-linear fits of the experimental data using this kinetic model confirm that it well represents the majority of observed duplex formation rates (Figure 2-13 below).

$$\ln(A) = a + b \cdot E_a \tag{10}$$

$$C \equiv \exp(a) \qquad T_c \equiv \frac{1}{k_b \cdot b} \tag{11}$$

$$k = \mathrm{C} \cdot \exp\left(\frac{E_a}{k_b}\left(\frac{1}{T_c} - \frac{1}{T}\right)\right) \tag{12}$$

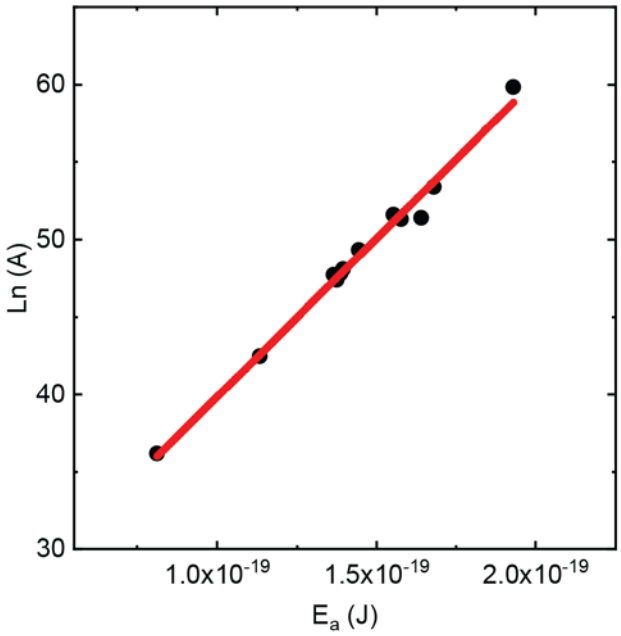


Figure 2-12 Observed correlation between the natural log of the Arrhenius pre-exponential (vertical axis) and the Arrhenius activation energy (horizontal axis). A linear fit to the data (red line) resulted in an R^2 value of 0.9889, an intercept of 19.4, and an intercept of 2.05 x 10^{20}.

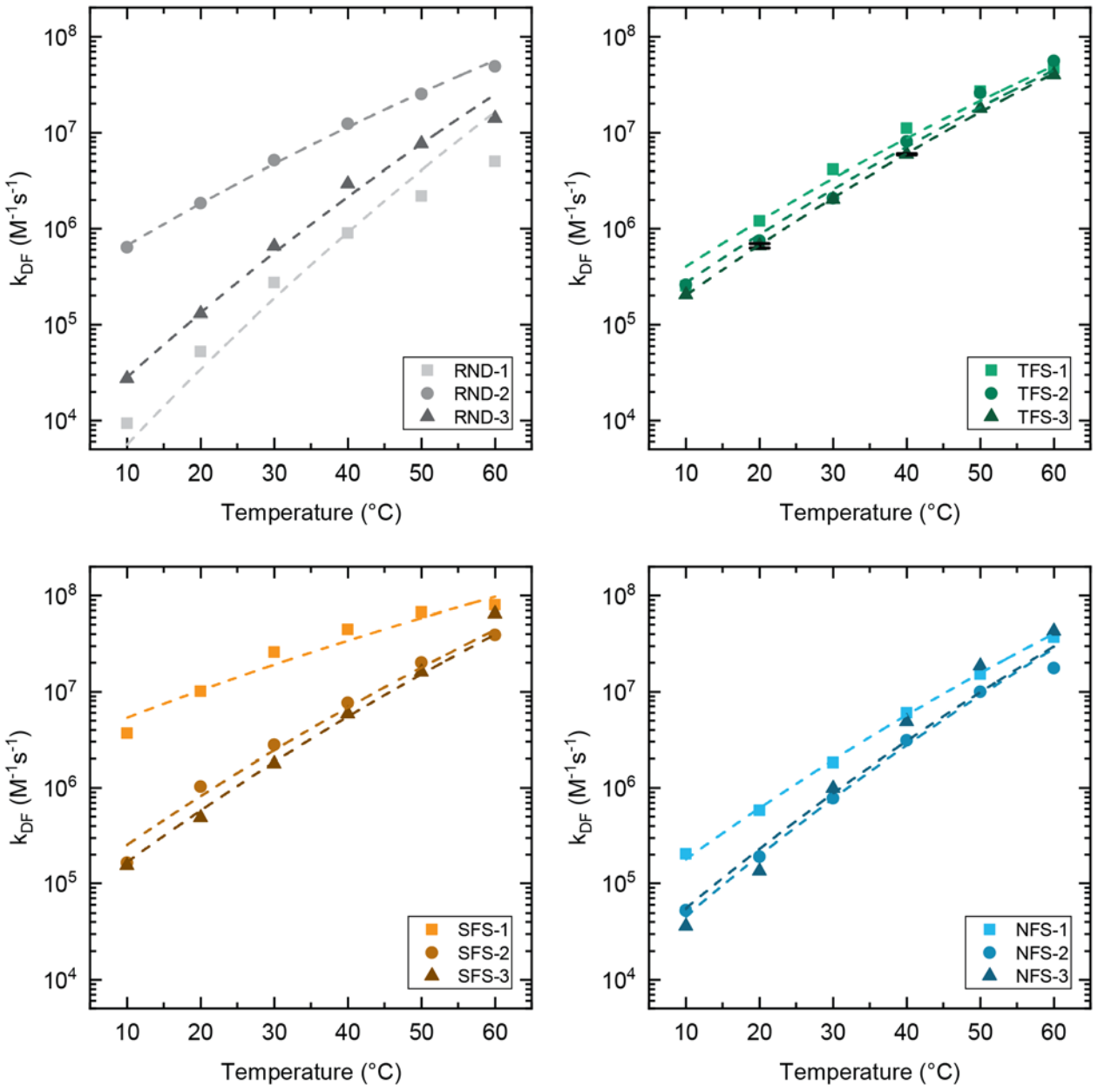


Figure 2-13 Experimentally determined duplex-formation rates (symbols) modeled using the empirically derived kinetic model reported in the text (dashed lines).

In the strand-displacement reaction, most implementations were found to possess rates exhibiting a parabolic dependence on temperature. Furthermore, systems were

found to only exhibit this parabolic behavior when they contain no intramolecular structures longer than 2 bp. Based on this observation, it was concluded that intramolecular structures lead to a change in the rate-limiting mechanism of the reaction leading to two distinct behaviors: (1) an approximately linear region at low temperatures with positive slope, and (2) an approximately linear region at high temperatures with negative slope. Both behaviors can be explained by a kinetic model where reactants form a stable intermediate which may then either proceed to reaction completion or dissociate. The mathematics of such a model are identical to the nucleation-and-zippering model of DNA duplex formation described in equations 8 and 9 above, albeit with varying physical interpretation of the relevant rate constants. In the case of the strand-displacement reaction, the stable intermediate is a three stranded complex and this complex proceeds to completion via the strand-displacement step. If the dissociation rate of the intermediate (k_{1r}) is much smaller than the rate at which the intermediate is converted into products (k_2), then k_{app} is approximately the duplex-formation rate of the toehold (k_{1f}). Based on the temperature-profile of the measured duplex-formation rates (Figure 2-9), it is reasonable that these rates may be Arrhenius with positive slope and thus explain the observed low-temperature behavior.

In addition, strand displacement rates (k_{SD}) were observed to converge at higher temperatures, and have negative slopes. This behavior can be explained based on the same kinetic model (Eq. 8, 9) if the rate of intermediate dissociation (k_{1r}) is much larger than the rate of intermediate conversion (k_2). This leads to apparent bimolecular rates (k_{SD}) which take the form of equation 13 below. Furthermore, modeling each of the three reaction rates k_{1f}, k_{1r} and k_2 as Arrhenius (eq. 7) results in an apparent bimolecular

reaction rate which is itself Arrhenius and possesses an activation energy ($E_{a,SD}$) equal to the difference in energies of the three barriers (eq. 14). In such a situation, a large energy barrier to intermediate dissociation ($E_{a,1r}$) may dominate the apparent energy barrier ($E_{a,SD}$) and lead to kinetics which depend almost exclusively on this value. This can be expected to result in rates which decrease as temperature increases and which depend strongly on toehold sequence, a variable held constant in these systems. The parabolic behavior can thus be described as a transition between the first behavior at low temperature and the second behavior at high temperature.

$$k_{SD} = \frac{k_{1f}}{k_{1r}} k_2 \qquad 13$$

$$\mathrm{E}_{a,SD} = \mathrm{E}_{a,1r} - \left(\mathrm{E}_{a,1f} + \mathrm{E}_{a,2}\right) \qquad 14$$

For both reactions studied, systems in the TFS design group exhibited the greatest kinetic reproducibility (Figure 2-14). For the duplex-formation reactions (Figure 2-14a), TFS-fit sequences were observed to possess temperature averaged M/MAD ratios of 19.4, corresponding with typical kinetic variations of ± 5%. For the strand-displacement reactions (Figure 2-14b) these devices exhibited temperature averaged M/MAD ratios of 9.1, corresponding to typical kinetic variations of ±11%. It is also evident that for temperatures near 20°C the duplex-formation and strand-displacement rates of TFS-fit systems are similar in value (approximately 1 x 10^6 $M^{-1}s^{-1}$). This fact may be of use to researchers looking to approximate the kinetics of DNA systems using a constant bimolecular rate.

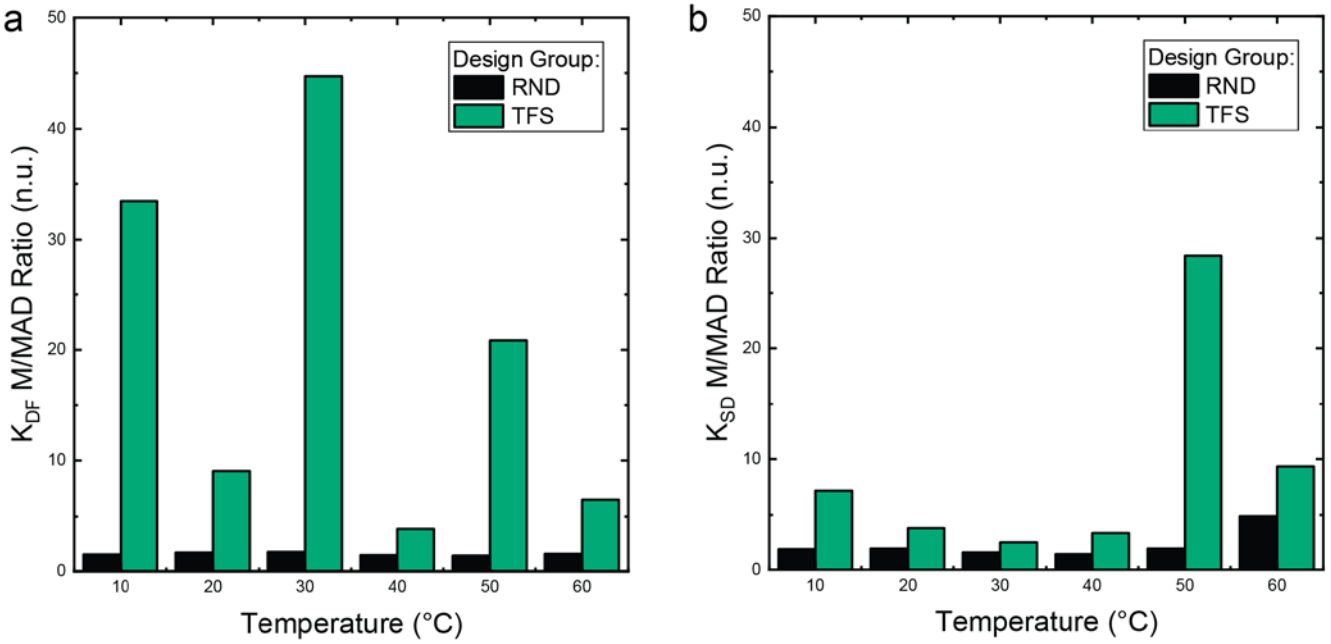


Figure 2-14 Kinetic reproducibility of the TFS-optimized implementations compared to the unoptimized RND implementations. (a) M/MAD ratios calculated from the duplex formation (k_{DF}) rates. (b) M/MAD ratios calculated from the strand displacement (k_{SD}) rates.

Section 2.5 – Further Discussion

DNA molecules have been previously shown to form numerous structures other than the A/T and G/C base pairs and resulting B-DNA double helix.[17] However, the results from this study indicate that the absence of intramolecular non-target B-DNA type structures results in DNA systems with highly reproducible kinetic rates. This surprising result suggests that although many alternative structures may form, B-DNA type structures are the primary contributor to kinetic variation. It is further observed that intramolecular structures as short as 3 bp exhibit a marked impact on reaction kinetics. Conversely, no experimental or theoretical evidence is observed linking small intermolecular structures to kinetic variation. The ability of large intermolecular structures to impact reaction rates has been well established[25,26,38,54,57,58,62-64], and is again

demonstrated in the strand displacement rates measured for the model system. However, SFS-fit systems selected from the Zhang *et al.* and Hata *et al.* datasets contain intermolecular complements from 3 to 7 bp in length yet remain kinetically uniform. This demonstrates that intermolecular structures up to this length can exist in DNA systems without substantially impacting reaction kinetics. As such, further study is necessary to establish under what conditions (i.e., length, location, and frequency) intermolecular structures will impact reaction kinetics.

Three properties were proposed quantifying the kinetic reproducibility of DNA systems: (1) SFS for quantifying intramolecular structures, (2) NFS for quantifying intermolecular structures, and (3) TFS which is a weighted linear combination of the first two. Of these three metrics, the SFS and TFS values which prioritize SFS were observed to explain the majority of kinetic variation. By analyzing experimental rate constants reported in the literature, it was shown that SFS-fit systems exhibit hybridization rates with M/MAD ratios of up to 13, a substantial improvement over the value of 1.7 observed for other sequences. This finding was confirmed by the creation and characterization of engineered TFS-fit systems. These sequences were observed to possess hybridization rate M/MAD ratios of 19.4, equivalent to variation of ± 5%. To date, the most accurate known model for predicting hybridization rates is the "6-factor" model derived by Zhang *et al.*.[36] Under specific experimental conditions, this model is capable of predicting hybridization rates within a factor of 3 for 91% of sequences. This level of accuracy translates to M/MAD ratios of ~2 or variations of ± 50%. This can be directly compared to the ± 5% observed for TFS-fit systems. As such, systems generated

using the new TFS criteria should therefore be expected to be substantially more reproducible than systems generated using current state of the art methods.

Evaluating sequence-set fitness using TFS has several key advantages relative to alternative methods. First, since TFS calculation is based solely on strand sequence without accounting for experimental conditions, sequence-sets which are TFS-fit are expected to perform similarly at a range of experimental conditions including temperature, buffer, and ion concentration. Second, calculating TFS does not require computation of thermodynamic parameters for the system making this method computationally efficient by comparison. Equivalently, this enables more potential systems to be considered in a fixed unit of time relative to thermodynamic approaches.

However, evaluating the fitness of DNA systems using TFS also has several key limitations. Foremost, it is clear that not all structures impact reaction kinetics equally. As such, TFS penalizes a number of systems which are kinetically-fit, yet contain non-problematic structures. Secondly, TFS penalizes systems based on the length of structures, this is based on the approximation that structure stability is based solely on length. For small structures, this approximation is not bad, however it degrades quickly as length scales. It is assumed that this will impact the effectiveness of TFS for systems which necessitate the inclusion of larger structures.

Section 2.6 – Conclusions

The kinetic variation observed in published reaction rates is strongly correlated with the presence of relatively small intramolecular B-DNA type structures. The custom-written SeqEvo software was used to generate new systems optimized to prevent these structures. By experimentally characterizing these systems, it was demonstrated that

engineering DNA systems to eliminate all non-target intramolecular B-DNA structures longer than 2 bp leads to kinetically uniform reaction rates. Engineering systems such that intramolecular structures larger than 2 bp are eliminated and intermolecular structures are otherwise minimized is recommended as a sequence-generation strategy. It is expected that this strategy will lead to systems with duplex formation rates varying by ± 2%, a marked improvement over the ± 50% which is the current state of the art.

CHAPTER THREE: AN ALGORITHM FOR GENERATING FIT SETS OF DNA OLIGONUCLEOTIDES

Knowing what to look for is not very useful if you don't know where to look. Similarly, having a fitness function is pointless without an appropriate search algorithm. For the reliable generation of DNA systems with uniform behaviors, two performance criteria are important: (1) the quality of the generated sequences is essential, and (2) it is necessary that the algorithm be efficient enough to produce a result in a reasonable amount of time. Here, an evolution-inspired algorithm for identifying fit systems is presented and studied.

Section 3.1 – The Remarkable Number of Potential Systems

For small systems, identifying fit sequences is not very difficult. Consider again a system composed of two fully complementary DNA strands (Figure 1-5). If these strands are each 6 bases long, then there are $4^6 = 4{,}096$ possible systems. As such, it would be relatively straightforward to use a computer and calculate the fitness of each possible system in order to identify the fittest system. In principle, one could even do it by hand. However, since the number of states for a given length L scales as 4^L, methods based on exhaustive calculation become unrealistic rather quickly. Indeed, for the simple 49 base strands studied in Figure 2-5, there are a staggering 3×10^{29} possible sequences. This number continues to grow exponentially as system complexity increases. For example, the 10x10x10 DNA-brick structure published by Ke *et al.* contains 7,824 variable bases, implying this design could be implemented by more than 10^{4710} possible DNA systems.

As mentioned previously, there are at least 12 software tools available for creating DNA systems. The tools which conduct *in silico* optimization utilize one of two approaches; either a bottom-up approach where systems are algorithmically created from fit sub-pieces,[40-42] or evolutionary approaches where quality is refined through iterative cycles.[15,50,52] Two strengths of evolutionary algorithms include their robustness to varying fitness landscapes and their ability to identify good solutions relatively quickly. Alternatively, a major strength of the bottom-up approaches includes their ability to create systems of exceptional quality. In order to ensure our design methodology is both widely applicable and effective, the decision was made to implement an evolutionary heuristic algorithm.

Section 3.2 – An Evolutionary Heuristic Algorithm

The proposed evolutionary heuristic algorithm is detailed in Figure 3-1 below. The search process is composed of five nested for-loops, each of which is controlled by a key parameter (Figure 3-1a). The five loops are described at an abstract level in Figure 3-1b, and a more detailed level in Figure 3-1c. The five parameters and their abbreviations are: *NL* (Number of Lineages), *CPL* (Cycles Per Lineage), *NMPC* (Number of Mothers Per Cycle), *GPC* (Generations Per Cycle), and *NDPM* (Number of Daughters Per Mother). The key process of system propagation and mutation are presented in Figure 3-2.

Algorithm execution can be described linearly to help understand the importance of the key parameters. Initially, the algorithm is provided a domain-level design and a sequence for each domain. Domains are declared as either variable, in which case they may be manipulated by the algorithm, or as fixed, in which case they will remain

unchanged by the algorithm. (As an example of this, consider the system presented in Figure 2-5 where the toehold is fixed while the other domain is variable.) From this system, a predetermined number (algorithm parameter NL) of clones are created. These systems are referred to as *lineage mothers* since each spawn their own independent evolutionary tree of descendant systems. All but one of the lineage mothers are mutated using type-1 or "large" mutations (Figure 3-2b), which result in the total randomization of all bases in each system. Each *lineage mother* repeats a structured evolutionary cycle a predetermined number (CPL) of times. Each cycle is composed of the following process: (1) A predetermined number (NMPC) of clones of the current *lineage mother* are created and termed *cycle mothers*, (2) All but one *cycle mothers* receives a Type 2 (medium) mutation where a random sequence of bases is relocated within a domain, (3) A predetermined number (NDPM) of clones of each *cycle mother* are created and termed *cycle daughters*, (4) All *cycle daughters* receive Type 3 (small) mutations where two random bases within a domain are swapped, (5) The fittest daughter replaces/becomes the cycle mother, (6) The process is repeated from step 3 a predetermined number (GPC) of times, and (7) The fittest cycle mother replaces/becomes the lineage mother. In our description of the algorithm, this seven-step process is referred to as an evolutionary cycle, despite the fact it is both composed of and part of larger cycles. After each lineage has undergone the predetermined number of cycles, the lineage mothers are compared and the fittest is identified. The resulting fittest system is returned as an output and the algorithm ends.

- For each of **NL** lineages, create a *lineage mother* system
  - For each *lineage mother*, run **CPL** cycles
    - For each cycle, create **NMPC** *cycle mother systems*
      - For each *cycle mother*, run **GPC** generations
        - For each generation, create **NDPM** *daughters*
        - At the end of each generation, the fittest *daughter* replaces/becomes the *cycle mother*
    - At the end of each cycle, the fittest *cycle mother* becomes the lineage mother
- At the end of all cycles, the fittest *lineage mother* becomes the victor

Parameter	**Abbreviation**
Number-of-Lineages	**NL**
Cycles-Per-Lineage	**CPL**
Number-of-Mothers-Per-Cycle	**NMPC**
Generations-Per-Cycle	**GPC**
Number-of-Daughters-Per-Mother	**NDPM**

- Algorithm Start
- gen0 = input system
- For (i = 0; i < **NL**; i = i +1)
  - If i equals 0, then lineage mother = new system (gen0, no mutation)
  - If i does not equal 0, then lineage mother = new system (gen0, large mutation)
  - For (j = 0; j < **CPL**; j = j +1)
    - For (k = 0; k < **NMPC**; k = k +1)
      - If k equals 0, then cycle mother = new system (lineage mother, no mutation)
      - If k does not equal 0, then cycle mother = new system (lineage mother, medium mutation)
      - For (l = 0; l < **GPC**; l = l +1)
        - For (m = 0; m < **NDPM**; m = m +1)
          - daughter = new system (cycle mother, small mutation)
        - End For
        - For each daughter
          - If score (daughter) ≤ score (cycle mother), then cycle mother = new system (daughter, no mutation)
      - End For
      - For each cycle mother
        - If score (cycle mother) ≤ score (lineage mother), then lineage mother = new system (cycle mother, no mutation)
    - End For
  - End For
- End For
- For each lineage mother
  - If score (lineage mother) < score (gen0), then gen0 = new system (lineage mother, no mutation)
- Return gen0 as the output
- End Algorithm

Figure 3-1 Pseudocode and key parameters describing the evolutionary algorithm. (a) High-level pseudocode illustrating the structure of the nested for loops and the naming conventions. (b) Parameters controlling the structure of the search. All parameters are given a positive integer value at runtime. (c) More detailed pseudocode further illustrating the search process.

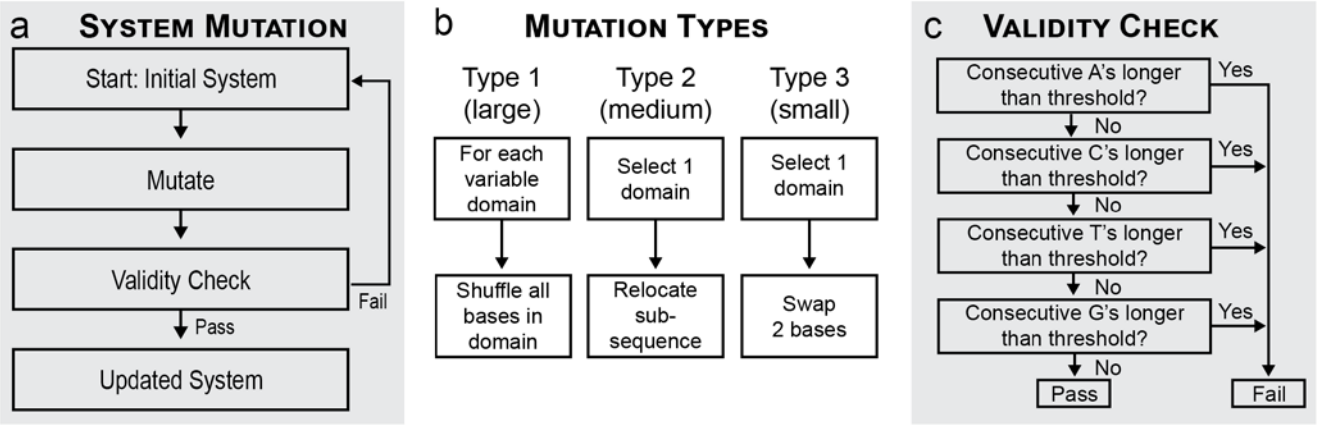


Figure 3-2 The algorithm utilizes a clone-then-mutate approach to generate new sequences. (a) Diagram illustrating the process for mutating a system's sequences. (b) The three types of mutations. (c) Diagram illustrating how valid/invalid systems are identified.

A set of example parameters (Figure 3-3a) and a visualization of the resulting search (Figure 3-3b) are presented below. The search starts with a single system (generation 0). From generation 0 three *lineage mothers* are created (generation 1 in the leftmost column), one of which is identical to the original (illustrated by the fact it is directly below generation 0). The other two *lineage mothers* have had their sequences randomized and have a high degree of uniqueness relative to generation 0. From each of these *lineage mothers*, three *cycle mothers* are created (generation 2). This step is also the start of the first evolutionary cycle (cycle #1 in the left column). Each set of *cycle mothers* contains one which is identical to the *lineage mother* and two which have had medium mutations applied. For each of the *cycle mothers*, two *daughters* are created (generation 3). These *daughters* receive minor mutations. The fitness of each *daughter* is calculated and compared to the *cycle mother*. If any *daughter* is more fit than the *cycle mother*, it becomes/replaces the *cycle mother*. Based on the GPC value of 3, *daughters* are generated and selected two more times (generations 4 and 5). At this point the first

cycle is completed, and the fittest *cycle mother* replaces the *lineage mother*. This new *lineage mother* begins a new cycle by creating three new *cycle mothers* (generation 6, start of cycle 2). The cycle process is then repeated (generations 6-9 are repeats of 2-5). Based on the CPL value of two, at the end of the second cycle (generation 9) the fittest of the *lineage mothers* is selected as the victor and is returned to the user.

a

PARAMETER	VALUE BELOW
NL	3
CPL	2
NMPC	3
NDPM	2
GPC	2

b

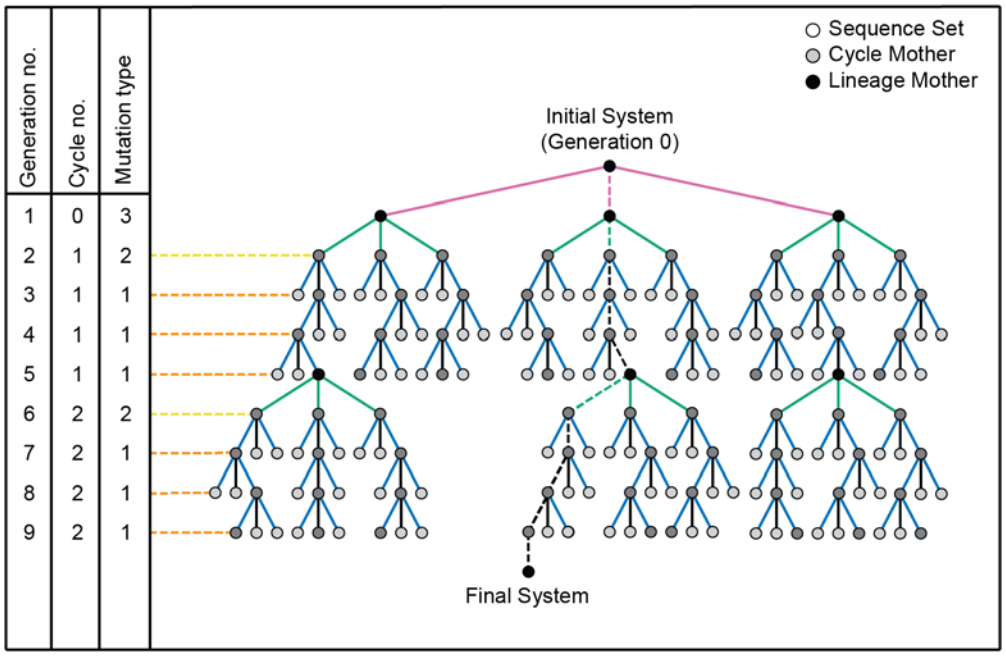


Figure 3-3 Example shape of the search for fit systems resulting from algorithm execution. (a) Example values for the five key parameters controlling the algorithm: Number-of-Lineages (NL), Cycles-Per-Lineage (CPL), Number-of-Mothers-Per-Cylce (NMPC), Number-of-Daughters-Per-Cycle (NDPC), and Generations-Per-Cycle (GPC). (b) Visual depiction of search progression for the example parameter values. Sequence uniqueness (horizontal axis) as a function of time (vertical axis, increasing downward).

Section 3.3 – Is the Algorithm Efficient Enough?

To study algorithm efficiency, the SeqEvo software was created which combined the evolutionary algorithm with the TFS fitness function. This software accepts as input a system's design, initial sequences, and a file containing the algorithm parameters to be used. A set of values for the five key parameters determines the shape of the evolutionary search and is referred to as a set of parameters or parameter-set. The relationship between search efficiency and parameter values was studied by systematically varying the parameter-set while keeping the other inputs fixed.

Method and Results

Systems were again defined as a set of DNA strands with declared sequence. Sequence quality was defined as either good (does not contain any non-target structures longer than 2 bp) or bad (contains non-target structures longer than 2 bp). Two strands forming a single 34 bp B-DNA duplex was chosen as a model system design. It has been previously demonstrated that this task is possible and that this is the largest such duplex which can be generated without introducing a 3 bp non-target structure.[41,42] The identification of a "good" system for this design was identified as a suitable design challenge for evaluating parameter-set effectiveness. This design challenge is a theoretically achievable result intended to validate that the software is operating properly.

Algorithm efficiency E was defined as 1/N, where N is the total number of systems which were considered prior to reporting a valid solution to the design problem. SeqEvo reports the score of each successive generation, making the calculation of this efficiency straightforward. Since algorithm performance is limited by the time necessary

to score each system, this quantity is expected to be proportional to the total computational resources consumed by both the algorithm and the software.

Global Efficiency Search

Using trial and error, an initial parameter-set capable of solving the design challenge was identified (Parameter set 0 in Table 3-1 below). This parameter set could reliably solve the design problem after considering ~256,000 systems, and consequently the search for efficient parameter sets was narrowed to the finite set of parameters-sets which considered up to ~2,560,000 systems. Starting with the minimal parameters of CPL=1, GPC=1, NDPM=1, NL=8, and NMPC=1, parameter-sets representing equal investment of 2,560,000 systems into 1, 2, 3, 4 or all 5 of the parameters were generated (parameter sets #1-31 in Table 3-1). The resulting 31 sets of parameters are expected to provide a course-grained sampling of parameter space. The SeqEvo software was run 81 times using each parameter set, and the resulting efficiencies calculated.

Table 3-1 Parameter values for each parameter-set used in the "Global" sampling of algorithm efficiencies.

	Parameter Values				
Parameter-Set	**CPL**	**GPC**	**NDPM**	**NL**	**NMPC**
0	1000	1	1	64	1
1	1	1	1	8	160000
2	1	1	1	512000	1
3	1	1	160000	8	1
4	1	160000	1	8	1
5	80000	1	1	8	1
6	1	1	1	1131	1131
7	1	1	565	8	565
8	1	565	1	8	565
9	400	1	1	8	400
10	1	1	1131	1131	1
11	1	1131	1	1131	1
12	800	1	1	800	1
13	1	400	400	8	1
14	400	1	400	8	1
15	400	400	1	8	1
16	1	1	137	137	137
17	1	137	1	137	137
18	109	1	1	109	109
19	1	69	69	8	69
20	68	1	68	8	68

	Parameter Values				
Parameter-Set	**CPL**	**GPC**	**NDPM**	**NL**	**NMPC**
21	68	68	1	8	68
22	1	109	109	109	1
23	109	1	109	109	1
24	109	109	1	109	1
25	55	55	55	8	1
26	1	40	40	40	40
27	40	1	40	40	40
28	40	40	1	40	40
29	24	24	24	8	24
30	34	34	34	34	1
31	19	19	19	19	19

Of the 31 parameter sets, 24 reliably identified a solution to the design challenge (> 80% success). A statistical summary of the observed efficiencies is reported in Figure 3-4 below, where data from the 7 ineffective parameters is omitted and replaced with an asterisk (*). This can be explained by the minimal values of CPL and GPC present in these parameter sets. Since these parameters control the number of iterative generations the algorithm undergoes, these types of searches are shallow in the sense that they consider a large number of random sequences, without allowing for iterative improvement of system quality.

For effective parameter sets, median efficiencies were observed to vary more than 2 orders of magnitude. The most efficient parameter set was observed to be parameter set

#5 (CPL = 80000, GPC = 1, NDPM = 1, NL = 8, NMPC = 1), which has been marked in orange in Figure 3-4. A typical design trial of the software using these parameters considered approximately 14,000 systems before solving the design challenge, whereas the next most efficient parameter set (#27 in Table 3-1) typically required about three times as many considerations (approximately 43,000 systems). For reference, there are $4^{34} \approx 3 \times 10^{20}$ possible sequences for the model system. This demonstrates that the algorithm is successfully able to identify high-quality sequences after considering only a small fraction of the possible systems.

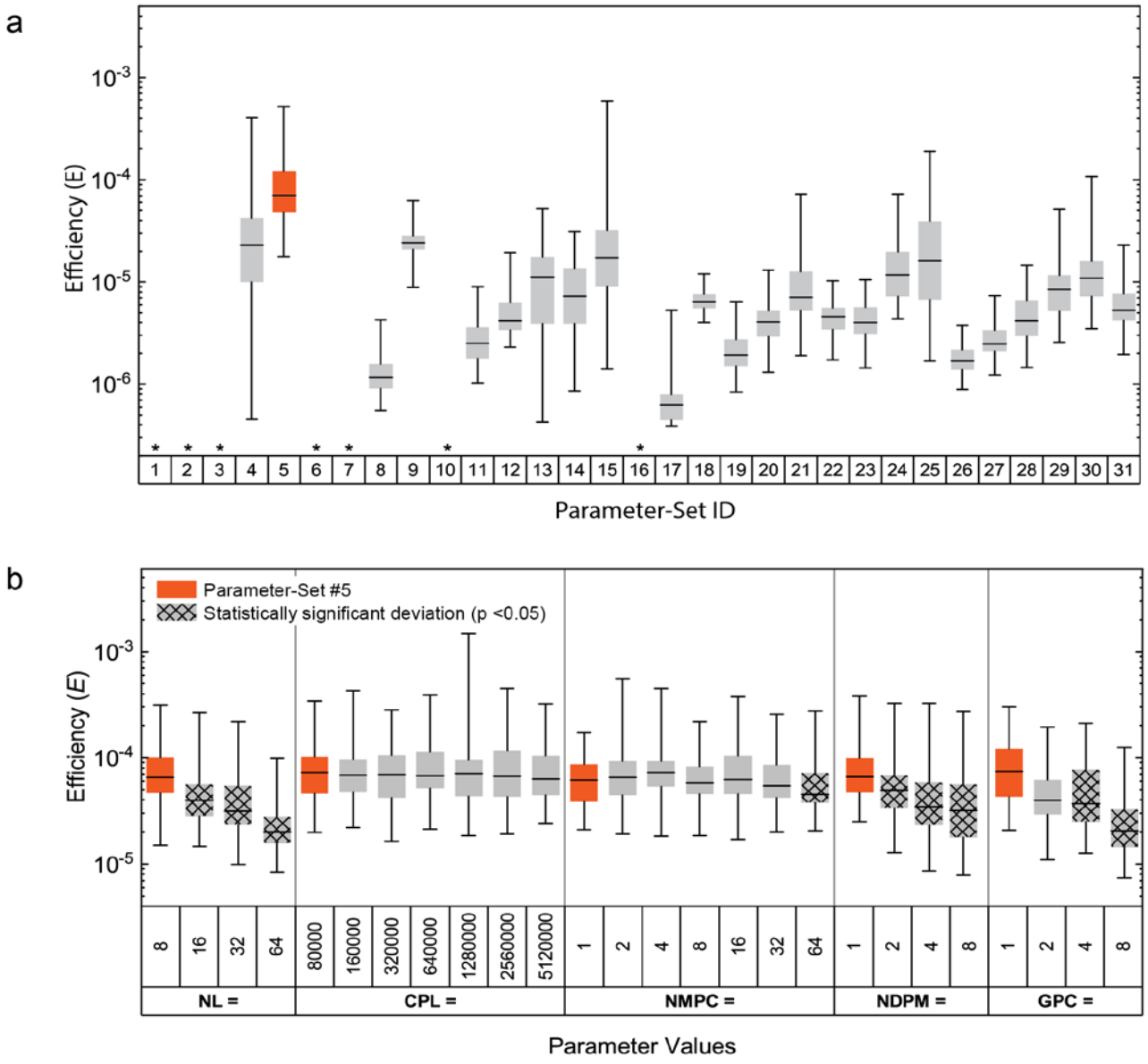


Figure 3-4 Efficiencies measured for varying parameter-sets. (a) The global search for efficient parameter sets. Efficiencies were measured for 31 parameter-sets spanning parameter space. For each parameter-set 81 independent design trials were conducted, and the efficiency was calculated for each. The observed efficiencies are summarized using a median line, box connecting the 25th and 75th percentiles, and bars connecting the min and max values. (b) The local search for efficient parameters. Parameter-set 5 (orange box) was identified as a highly efficient region in parameter-space. This region was investigated in greater detail by systematically varying each parameter while monitoring efficiency. For each new parameter set, 81 independent trials were again conducted. The 81 derived efficiencies are again summarized using a median line, a box connecting the 25th and 75th percentiles, and bars connecting the min and max values. Statistical significance was calculated using a two sample Kolmogorov-Smirnov test with p-value of 0.05.

Local Search. The area surrounding parameter-set #5 (Orange in Figure 3-4a) was further explored to confirm a local maximum in efficiency (Figure 3-4b). Parameter-set #5 resides on the boundary of parameter space, with a CPL value of 80,000 and all other parameters at their minimal value. Each of the five key parameters were systematically increased while holding all other parameters fixed to the values from parameter-set #5. The ranges over which the parameters were varied are specified on the horizontal axes in Figure 3-4b. The efficiency of these additional 21 algorithms were determined using a similar 81 trials-per-parameter-set approach. A statistically significant (Kolmogorov-Smirnov test with P-value < 0.05) decrease in efficiency was observed immediately for the NL, NDPM, and GPC parameters, indicating that increasing these parameters has an adverse effect on algorithm performance and should be avoided. Efficiency appears to be relatively stable for NMPC values up to 32, above which a decrease in efficiency was resolvable. No variation in efficiency was observed for the CPL parameter, which was anticipated based on the fact that CPL controls algorithm duration but has minimal effect on the structure of the evolutionary search. These results suggest that the region of high efficiency encompasses parameter sets with the following parameter values; NL of 8, NDPM of 1, GPC of 1, NMPC between (and including) 1 to 32, and no observed limitation on CPL. It is suggested that parameters of NL = 8, NDPM =1, GPC =1, and NMPC = 2 be used as default values, and that CPL be tuned depending on the algorithm runtime/system quality desired.

Discussion

The efficiency of the evolutionary algorithm was investigated and tuned using a model system. For the 34 bp duplex used as a model system design, the algorithm was

most efficient when the NL, NDPM, and GPC parameters were minimized. Increasing the CPL and NMPC parameters did not appear to have a strong effect on efficiency. The most efficient parameter-sets were able to reliably identify a high-quality design after considering approximately 14,000 potential systems. This represents a small fraction of the 3×10^{20} systems possible, indicating that the algorithm is able to identify fit systems with efficiency appropriate for this application.

Section 3.4 – How Effective is the Algorithm Compared to Other Software?

In order to confirm that the algorithm is functioning properly and that the resulting quality is an improvement over state-of-the-art, the SeqEvo software was benchmarked against several alternative tools.

Methods and Results

Two strands forming a single 35 bp DNA duplex were identified as an appropriate model system design. This is known to be the smallest duplex which requires the introduction of a three bp long non-target structure,[41,42] and is thus expected to highlight performance differences between the design methods. For each method, default parameters were applied in three independent trials. The interference profiles of the resulting eighteen devices were calculated using the DevPro software. New systems were generated using one of eight methods: two different TFS weighting factors, one of five alternative state-of-the-art sequence-generation methods, or random sequence assignment. The eight methods included: (1) the SeqEvo software utilizing scoring weights of $C_1 = 1$ and $C_2 = 1$, (2) the SeqEvo software utilizing scoring weights of $C_1 = 10{,}000$ and $C_2 = 1$, (3) the Domain Design software,[15] (4) the DNASequenceGenerator software,[41] (5) the Exhaustive Generation of Nucleic Acid Sequence (EGNAS)

software,[42] (6) the Nucleic Acid Package (NUPACK) software,[50] (7) the Uniquimer3D software,[48] and (8) random sequence assignment via a pseudo-random number generator. The median interference profile (as judged by the NFS metric) was selected for each design method and is reported in Table 3-2 below.

Table 3-2 Typical non-target structures present in 35 bp duplexes generated using several publicly-available software tools.

	Non-Target Structure Type, Length (bp), and Count (№)									
	Intramolecular Length			Intermolecular Length						
Sequence Source	**4**	**3**	**2**	**8**	**7**	**6**	**5**	**4**	**3**	**2**
Random Sequences	4	16	64	0	0	0	10	50	202	250
SeqEvo (TFS 10,000:1)	0	0	0	0	0	0	6	36	90	176
SeqEvo (TFS 1:1)	0	0	34	0	0	0	0	0	2	154
Domain Design[15]	0	8	54	0	0	0	2	14	36	218
EGNAS[42]	0	8	58	0	0	0	0	0	30	204
DSG[41]	0	10	66	0	0	0	0	0	36	206
Uniquimer3D[48]	4	26	86	2	4	8	12	22	72	250
NUPACK[50,52]	6	22	70	0	0	2	6	30	76	232

Systems generated using randomly seeded sequences were observed to regularly contain intramolecular and intermolecular structures of 4 and 5 bp, respectively. The sequences containing the fewest intramolecular non-target structures resulted from application of the SeqEvo software (C_1 = 10,000 and C_2 = 1 scoring weights) and contained no such structures. These scoring parameters also resulted in the elimination of all intermolecular structures longer than 5 bp. The sequences containing the fewest intermolecular structures were also generated using the SeqEvo software (C_1 = 1, C_2 = 1

scoring weights) and contained no such structures longer than 3 bp. These scoring parameters also resulted in the elimination of all intermolecular structures longer than 2 bp.

Discussion

SeqEvo was demonstrated to outperform the alternative sequence generation methods for the model system. This is a positive yet expected result for two reasons: (1) SeqEvo is the only method which is specifically optimizing for the performance criteria, and (2) SeqEvo's algorithm was engineered to do precisely this.

SeqEvo's ability to outperform the other software is informative. First, this demonstrates that the algorithm, fitness function, and software tool are all operating properly. Second, it suggests that no currently available design software is eliminating intramolecular structures as effectively as SeqEvo. Finally, one major advantage of SeqEvo is expected to be its robust ability to generate high-quality sequences for more complex systems. In order to accommodate the limited range of designs certain programs could generate, it was necessary to limit the model system to a single DNA duplex. SeqEvo outperforming the other methods in this model system is a good indicator, and one should expect the performance difference to be more pronounced for larger and more complex systems.

Section 3.5 Conclusions

A new heuristic evolutionary algorithm for robustly generating fit sets of DNA sequences was proposed. This algorithm utilizes staged mutation-selection cycles to systematically identify systems with improved fitness scores. The pairing of this evolutionary algorithm with the TFS fitness function enabled creation of the SeqEvo

software tool. This software was then used to tune algorithm parameters and evaluate the efficiency of the algorithm. A model system (34 bp duplex) was used to identify that certain parameter-sets are much more efficient than others. This set of parameters is suggested as default values for the SeqEvo software.

The performance of the fitness score and algorithm were evaluated by benchmarking against alternative state-of-the-art methods using another model system (35 bp duplex). Performance was based on system-quality and quantified by counting the non-target structures in the generated systems. SeqEvo was observed to outperform all studied alternative methods in resulting device quality. It was postulated that the performance difference would be more pronounced in more complex systems.

CHAPTER FOUR: SOFTWARE IMPLEMENTATION STRATEGIES

In principle, TFS is a lightweight fitness function, the heuristic algorithm is efficient, and their combination should be a highly effective new design tool. However, in order to deliver on this potential, proper software implementation is key. As such, care was taken to make the software both widely applicable and highly tunable.

Section 4.1 – Software Architecture Briefly Explained

Two software tools were created. The first tool, named Device Profiler and abbreviated DevPro, is intended to automate the calculation of non-target structures in a given system. The second tool, named Sequence Evolver and abbreviated SeqEvo, is intended to generate fit sequences for a target DNA system. Both software tools were written in the Java language for deployment across all prominent computing platforms (including Windows, MacOS, and Linux). The two tools share a large portion of their code. Specifically, DevPro is built around a modular piece of code referred to as the scoring module. SeqEvo is built around both this scoring module, and an additional module referred to as the heuristic module. The current version of DevPro and SeqEvo contain about 9,000 lines of code organized into 12 files. The source code of both programs has been made available in a GitHub repository (https://github.com/MTobiason/Sequence-Analysis). Care was taken to make the code useful in both multi-thread and multi-node situations. As a result, the code has been utilized successfully on machines ranging from a typical laptop computer, to Boise State's R2 High-Performance-Computing (HPC) cluster (DOI: 10.18122/B2S41H).

Section 4.2 – Strategies for Improving Software Efficiency

Efficiently Scoring Systems

Most of the computation required to calculate TFS is incurred while calculating the NFS term. Calculation of NFS requires the consideration of every possible base-pair for every possible strand combination in the system. For systems with more than one strand, the number of possible two-strand combinations is $n(n+1)/2$ (where n is the number of strands in the system). Each of these strand combinations have $(L_1 + L_2 - 1)$ possible strand alignments (where L_1 and L_2 are the number of bases in each strand), and each alignment must be checked for complementary sequences. As such, even simple systems possess many strand alignments which must be evaluated. As an example, consider again the model system presented in Figure 1-6. For this system, the design is two strands which form a complementary duplex. The strands are composed of only the alpha domain and its binding complement, both of which are 6 bases in length. The process used by the scoring module to analyze this system is shown in Figure 4-1 below. The module begins with the domain sequences (left). These are mapped onto the domain-level design to create strand sequences. For each combination of strands, the possible strand alignments are then exhaustively calculated. In figure 4-1, only the alignments resulting from Strand-1/Strand-1 interactions are displayed, and the "overlap" region which must be read is highlighted in bold. For this system there are 3 strand combinations, 33 possible strand alignments, and 108 possible base pairs which need to be considered. For a similar system composed of 49 base-pair duplexes, there are 3 strand combinations, 291 strand alignments, and 7,203 possible base pairs which need to be considered.

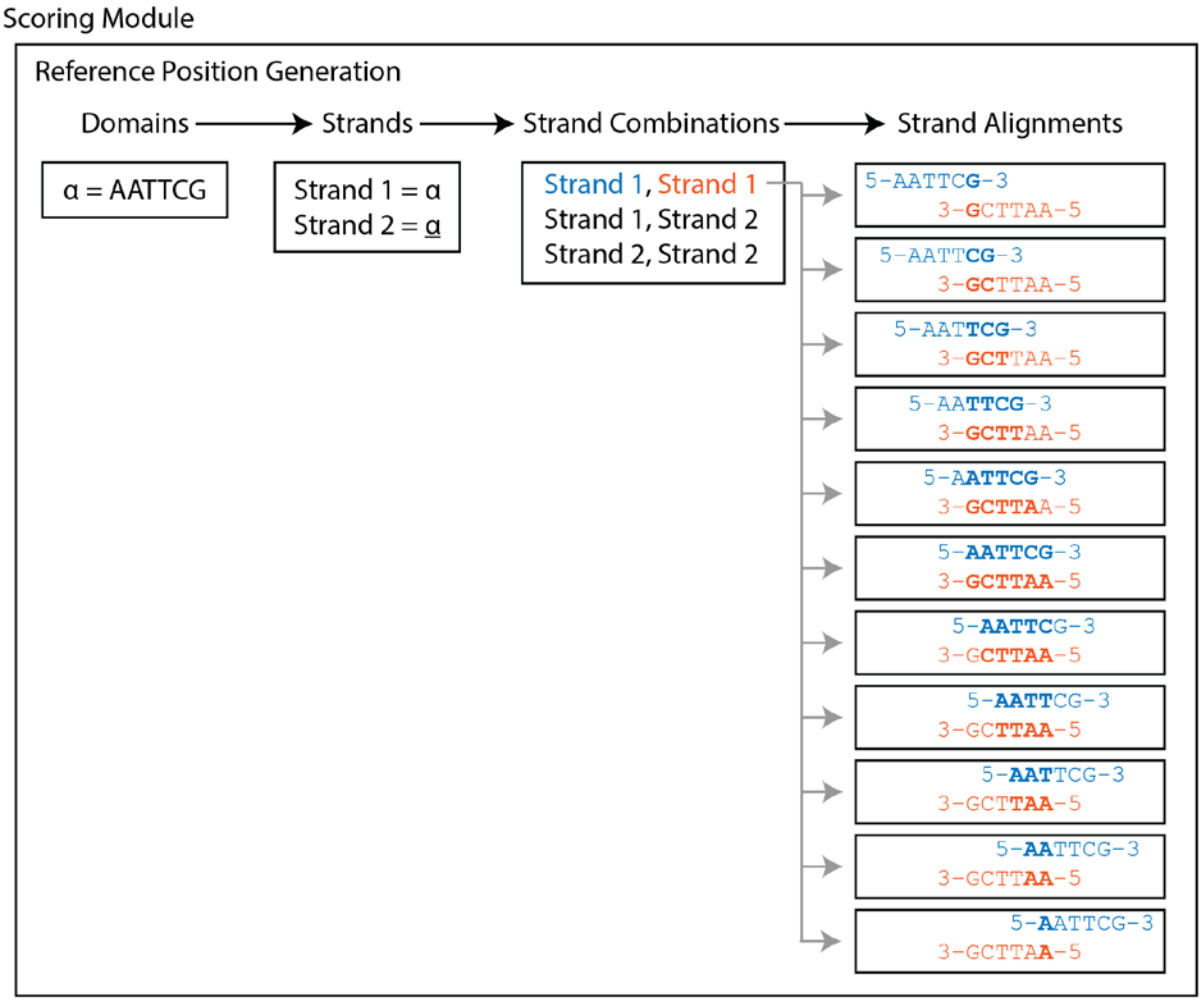


Figure 4-1 Illustration of how the scoring module calculates the strand alignments for a given system. Only the alignments for the Strand 1/Strand 1 combination are shown.

In the scoring module, the strand alignments for a given design are calculated only once. Every subsequent time a system is evaluated, the same alignments are used, but they are passed new sequences to evaluate. The process the scoring module uses to calculate the score is presented in Figure 4-2 below. For each given strand alignment (left). A set of structures is generated. Based on this set of structures, a structural profile listing the number of complements of a given length can then be generated. A score for each alignment can then be calculated. Finally, summing up the scores of each alignment

yields the total score (NFS in this case). SFS is calculated in a similar fashion to NFS, but the “strand combinations” are replaced simply by the list of strands, and the overlap region for each strand alignment is limited to only those which may form from a single molecule.

In order to calculate a detailed structural report, the DevPro software makes use of the full process described in Figures 4-1 and 4-2. In contrast, SeqEvo only needs to know the total score as quickly as possible. Therefore, two key shortcuts were introduced into the process. First, when the scoring module is asked directly for the score of a given alignment, it reads through the alignments and calculates the score without recording either the structure sequence or the structural profile. Second, the scoring of an entire system is only done once for newly generated systems. Anytime a system is generated via a mutation, the alignments impacted by this mutation are immediately rescored. The difference in scores between the alignments before and after mutation is removed from the previous TFS to calculate the new score.

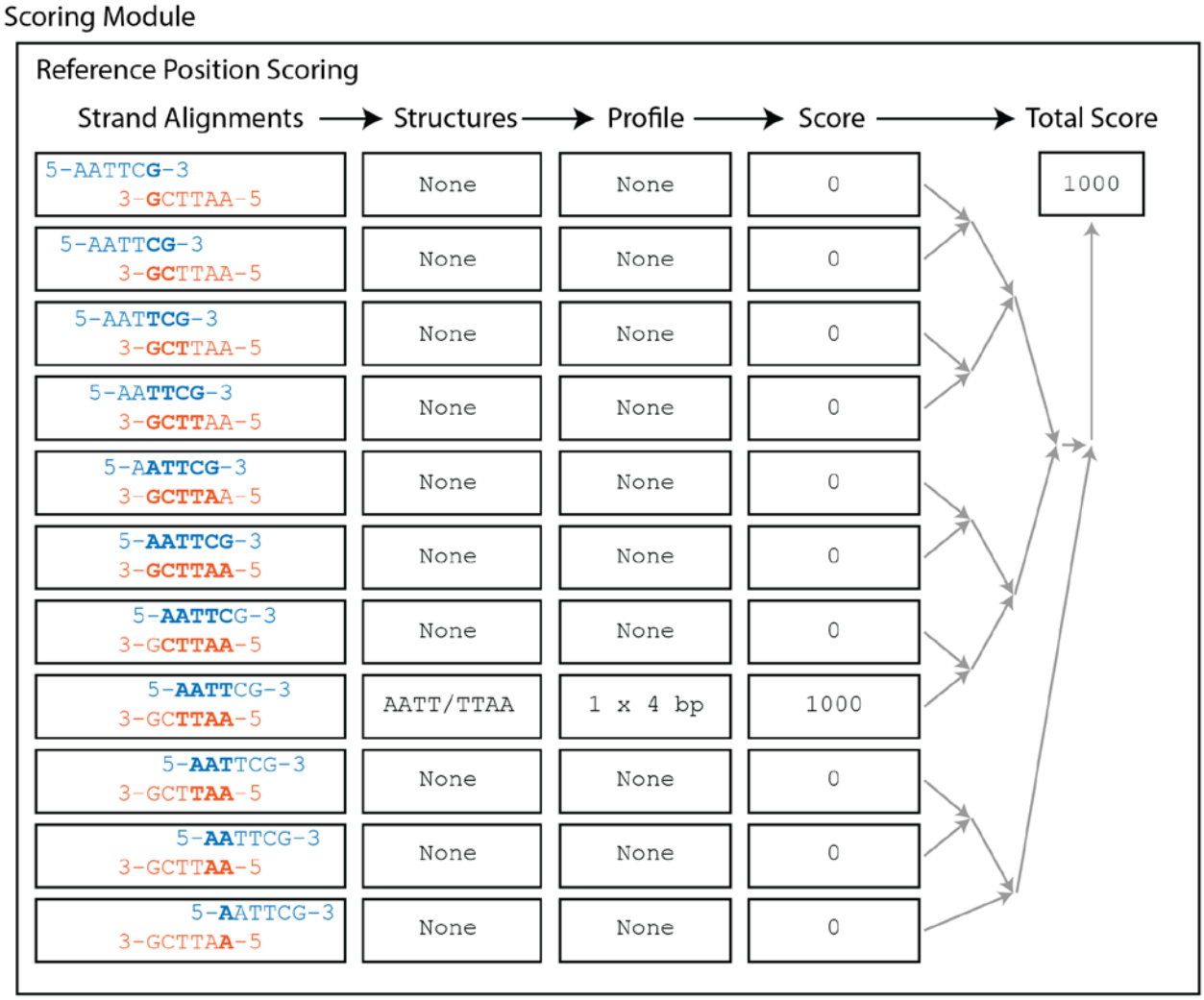


Figure 4-2 Illustration of how the scoring module calculates fitness scores for each strand alignment.

Minimizing Systems in Memory

In order to execute the evolutionary algorithm, many unique systems must be considered. To illustrate this, consider again the example search structure introduced in Figure 3-3 and reproduced in Figure 4-3 below. This relatively modest search considers a total of 122 systems with unique sequences. This number scales when increasing any of the algorithm parameters. In our study of algorithm efficiency, we observed that the consideration of at least 14,000 systems was necessary to solve even a modest design

problem (Figure 3-4). In order to minimize the number of systems kept in memory at any one point, an aggressive recycling strategy was used. This strategy resulted in the creation of a single object in memory for each lineage mother, cycle mother, and cycle daughter, and an additional one object recording the initial system (Figure 4-3b). For the example search structure, this approach results in 40 systems being kept in memory at a given time, roughly one third of the total systems considered. Importantly, the number of systems in memory does not scale with the CPL or GPC parameters which are typically used to control algorithm duration, so the one third ratio will approach zero as algorithm duration in increased using these parameters.

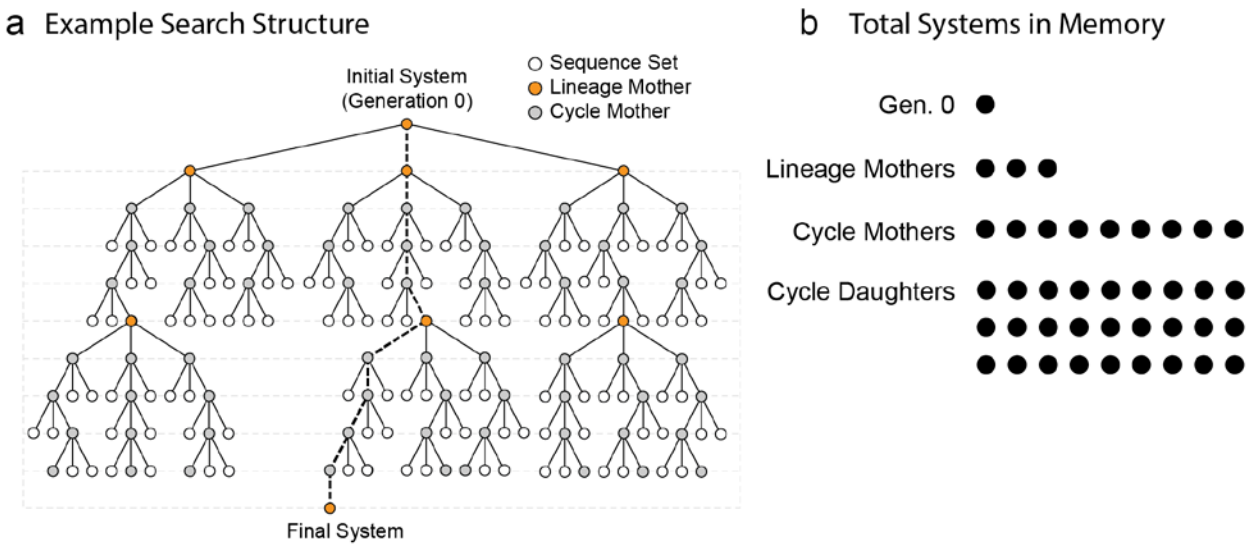


Figure 4-3 (a) Illustration of an example search structure. (b) The number of device objects which are kept in memory for the given search structure.

Section 4.4 – Can the Software Improve Published Systems?

To evaluate SeqEvo's ability to handle both complex and or large systems, several novel systems with published sequences were re-implemented using the software.

Methods and Results

New sequences for four systems, each created using a unique architecture, were generated using the SeqEvo software. For each system, the published domain-level design and sequences were identified and converted to SeqEvo input. Fixed design features such as G/C clamps were preserved. Software parameters were tuned for each device using a trial and error approach with a focus on eliminating all intramolecular events longer than 3 bp. Design trials were run on one of three platforms (a windows-based laptop computer or one of two Linux-based servers) to demonstrate the software's ability to be deployed on varying computer architectures. The four systems were: (1) the 10x10x10 DNA brick structure published by Ke *et al.*,[30] (2) the "four-input OR" seesaw-gate based network published by Qian *et al.*,[29] (3) the autocatalytic four-arm junction published by Kotani *et al.*,[65] (4) and the autocatalytic network published by Zhang *et al.*.[25]

The non-target structures present both before and after optimization are reported in Figures 4-4, 4-5, 4-6, and 4-7 for the four systems. As a result of the architecture of the SeqEvo program, the new sequences are re-arrangements of the bases present in the original design and are therefore expected to have similar thermodynamic stabilities for the target structures. For each of the system, the new sequences represent decreases in both the total number of interference events and the number of interference events of any given length and type. For the Brick system by Ke *et al.*, the original sequences were found to contain 4,062 interferences which are 3 bp or longer. The new sequences for this system contain only 67 intramolecular interferences which are 3 bp in length. For the Qian, Kotani, and Zhang systems, new sequences with no interferences 3 bp or longer

were generated, however the design of the Kotani system was observed to require substantially more computational resources to achieve this level of quality than the other two. This likely arises due to the increasing difficulty of finding intramolecular-interference-free sequences as strand length grows, and the inclusion of a 74-base strand in this system's design. For comparison, the Ke and Qian systems possess maximum strand lengths of 48 and 33 bases, respectively. The size of the largest non-target structures before/after optimization are reported for each of the four systems in Table 4-1 below. Most notably, SeqEvo was observed to reliably identify sequences with reduced non-target structure for each system.

The four systems were observed to require significantly different amounts of computational resources. The most modest resource consumption was for the Zhang *et al*. system. This system required about two minutes on a laptop computer to consider 800,000 potential systems and arrive at a quality solution. The most resource greedy of the designs was the Ke *et al.* system, which considered 48,000,000 potential systems in 6 hours utilizing four nodes of Boise State's R2 cluster (Dual intel Xeon E5-2680 v4 processors).

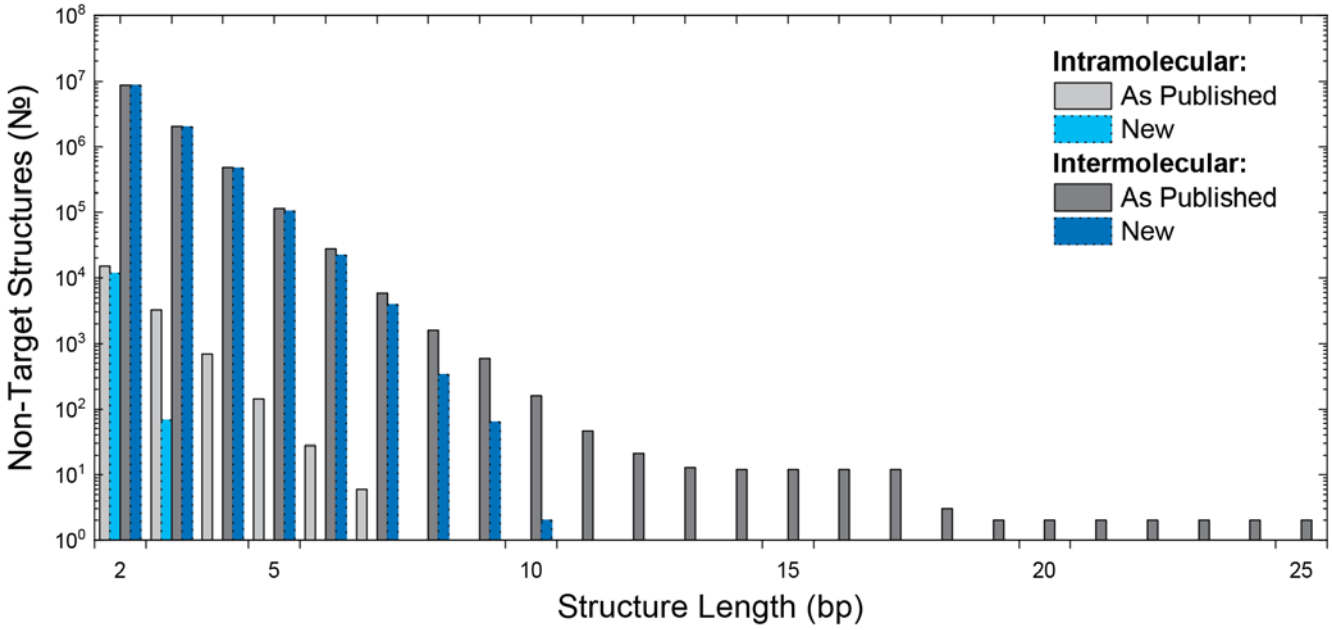


Figure 4-4 Profile of non-target structures in the 10x10x10 DNA-brick before (grey) and after (blue) sequence optimization.[30]

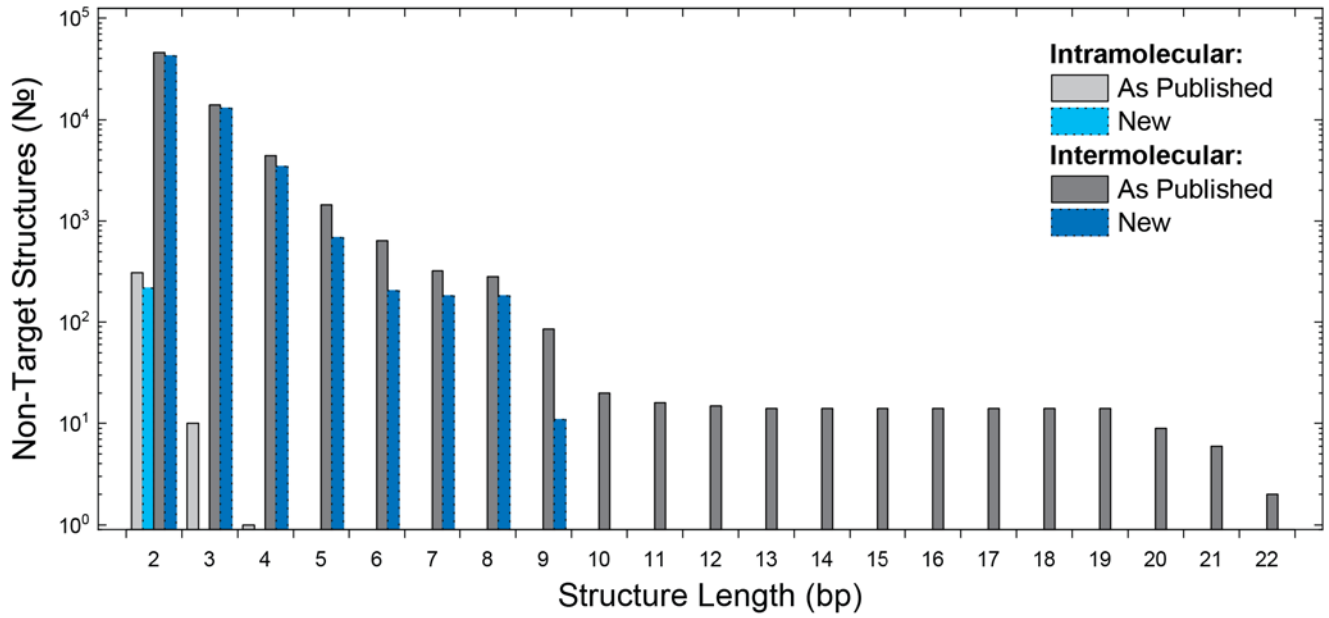


Figure 4-5 Profile of non-target structures in the "four-input or" seesaw-gate system before (grey) and after (blue) sequence optimization.29

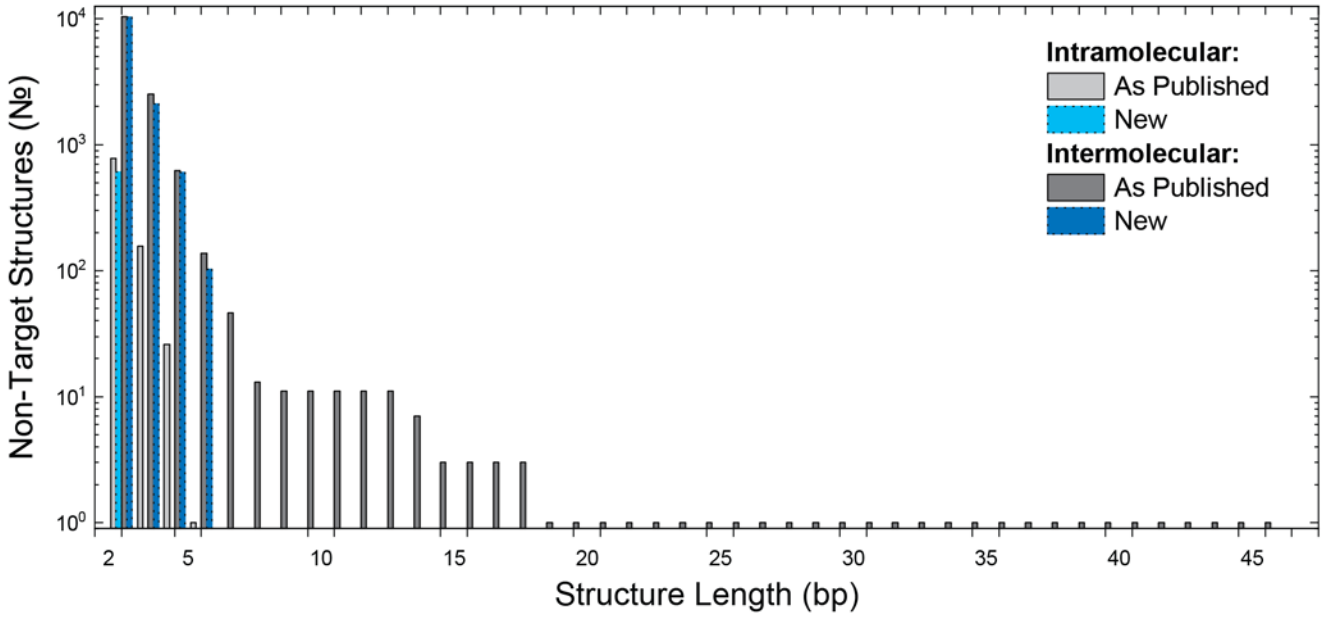


Figure 4-6 Profile of non-target structures in the autocatalytic four-arm junction system before (grey) and after (blue) sequence optimization.65

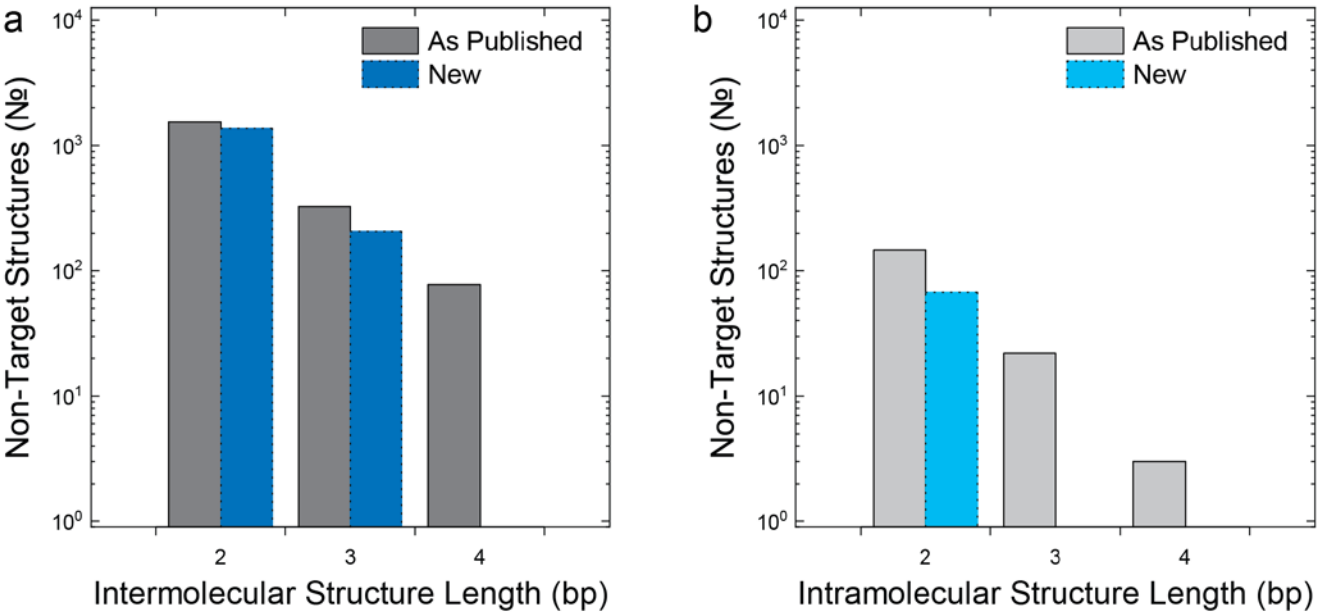


Figure 4-7 (a) Intramolecular and (b) intermolecular profiles of non-target structures in the autocatalytic system published by Zhang et al. before (grey) and after (blue) sequence optimization.25

Table 4-1 Select properties from the four re-engineered systems.

			Largest non-target structure (bp)	
Architecture	**№ Strands**	**Sequences**	**Inter-**	**Intra-**
Ke *et al.*[30]	517	As published	25	8
		TFS (10^8:1)	10	3
Qian *et al.*[29]	45	As published	22	4
		TFS (10^4:1)	9	2
Kotani *et al.*[65]	10	As published	45	5
		TFS (10^6:1)	6	2
Zhang *et al.*[25]	6	As published	4	4
		TFS (10^2:1)	3	2

Discussion

SeqEvo was observed to generate TFS-fit sequences for all four of the novel systems. For three of the four systems, all non-target intramolecular structures longer than 2 bp were eliminated. This level of quality implies exceptional kinetic reproducibility based on our prior evaluation of SFS effectiveness.

The largest system engineered was the 10x10x10 DNA-brick structure published by Ke *et al.* This structure is composed of 517 strands containing 17,248 bases and required six hours on four nodes of a HPC system. The current version of SeqEvo is expected to perform well for systems of this size or smaller.

Section 4.4 Conclusions

By re-engineering several established systems, it was demonstrated that the SeqEvo software can generate high quality sequences for a range of state-of-the-art systems. Furthermore, based on the software's implementation in the Java programming language, SeqEvo is expected to be usable on a variety of devices ranging from personal computers to HPC systems. The software's ability to accept domain-level designs is highly generalized and is capable of integrating with a variety of present and future design methods. Collectively, these factors should enable a wide range of researchers to use the program for the development of a wide range of DNA systems including but not limited to state-of-the-art dynamic and structural systems.

CHAPTER FIVE: ENGINEERING SYSTEMS WITH UNIFORM BEHAVIOR

A new method for generating DNA systems with uniform behavior is proposed based on three advancements to the state-of-the-art: (1) an improved criterion for identifying kinetically uniform devices *in silico*, (2) an improved algorithm for robustly identifying fit systems, and (3) two new software tools for automating the analysis and generation of systems.

Section 5.1 – Key Findings

Five key studies supporting the development of a new sequence generation method were conducted (Figure 5-1a below). First, recently observed kinetic variations were analyzed and found to be explainable by the presence of small intramolecular B-DNA type structures (Study 1). By quantifying such structures using the SFS and TFS properties, it was demonstrated that systems which contained fewer such structures exhibited kinetic behaviors which were significantly (P-value > 0.05) different than other systems. This finding was further strengthened by the creation and experimental characterization of TFS-fit systems (Study 2). Based on the results from these two studies, it was concluded that DNA systems which contain no intramolecular non-target structures longer than 2 bp should have duplex-formation rates varying by ± 5%. Based on the intuitive fact that large intermolecular structures will also impact device function, it was suggested that engineered systems also contain no non-target intermolecular structures larger than the target structures.

a

Study Number	Key Question
1	Do B-DNA structures explain kinetic variation?
2	Are TFS-fit sequences kinetically uniform?
3	Is the algorithm efficient enough?
4	How effective is the algorithm compared to other software?
5	Can the software improve published systems?

b

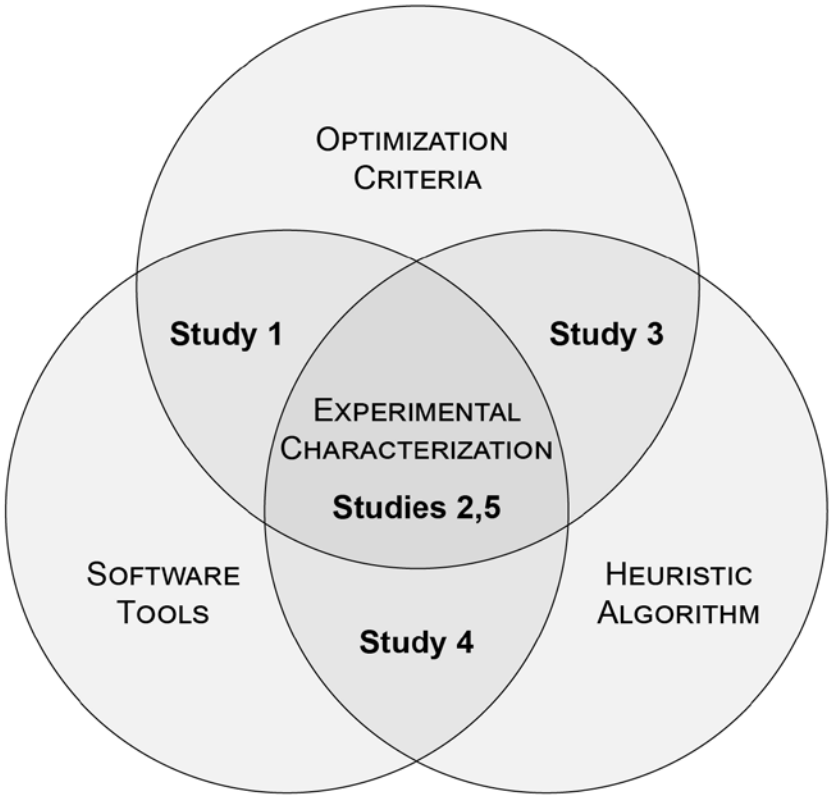


Figure 5-1 Key elements of this dissertation. (a) The five studies supporting creation of the design method. (b) Venn diagram illustrating the interconnected nature of the criteria, algorithm, and software. Studies have been generally associated with key areas to demonstrate their relative contributions to the dissertation.

To address the challenge of reliably identifying systems which are fit with respect to these criteria, an evolution-inspired heuristic algorithm was proposed. Both the efficiency (study 3) and the effectiveness (study 4) of this algorithm were characterized.

The algorithm was found to produce systems of higher quality than state-of-the-art tools, and to do so with appropriate efficiency.

Finally, to automate the process of characterizing existing systems and generating new systems, two software tools were developed. These tools are intended to be useful for researchers looking to develop a wide range of DNA systems, including those for both structural and/or dynamic applications. The ability of these tools to handle state-of-the-art systems was demonstrated by re-engineering several existing systems (study 5). The software was found to accommodate each of the existing systems, and resulted in high-quality implementations (no non-target intramolecular structures > 2 bp) for 3 of the 4 systems.

Section 5.2 – The method

Based on these studies, a formal method for creating high-quality systems using *in silico* sequence optimization was created. This method can be described as a process and is illustrated in Figure 5-2 below. The process begins with a domain-level design. At this stage, the system is described by a list of strand names each of which is associated with a series of binding domains and/or binding domain complements. Next, initial domain sequences are generated for each of the binding domains and each domain's sequence is declared as either variable or fixed. The design and domain sequences are then input into the SeqEvo software, and the system is optimized using the default parameter set. At the completion of this step, a set of strand names with associated sequences is generated. Next, the system output by SeqEvo is analyzed to verify device quality. The DevPro software is useful for this task. If the system contains intramolecular non-target structures larger than 2 bp or intermolecular structures larger than the target

structures, the software parameters are updated (such as increasing algorithm duration) and optimization is repeated. If the system contains none of these structures, it is deemed fit and can be experimentally characterized. If a high-quality system cannot be identified by tuning software parameters, it may be necessary to either refine the design of the system or relax the quality-requirements.

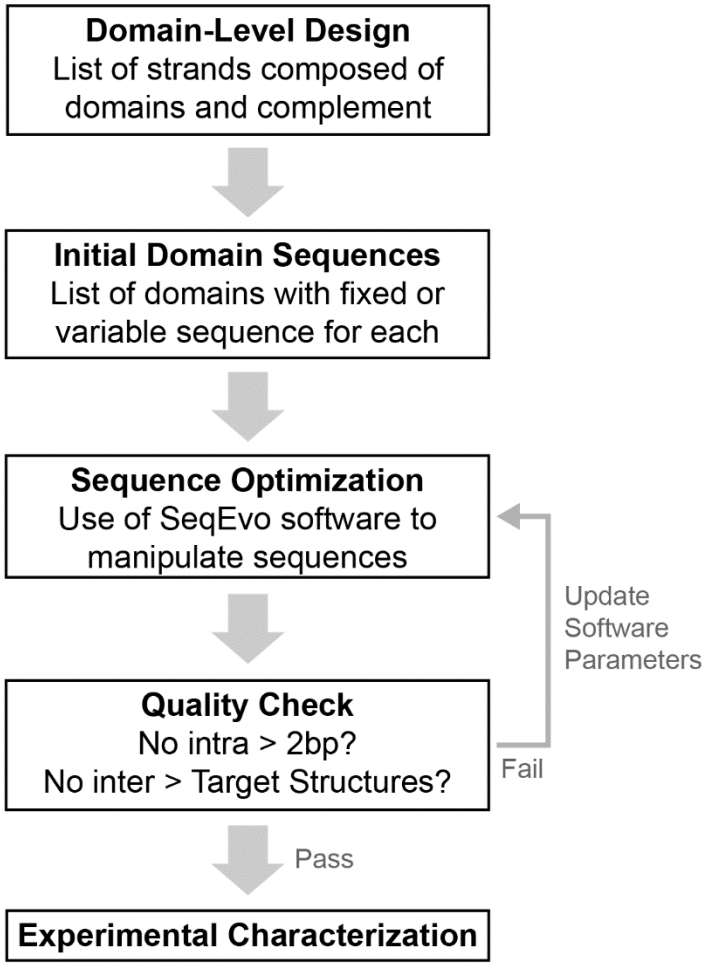


Figure 5-2 A process for creating kinetically uniform DNA devices utilizing *in silico* sequence optimization.

Section 5.3 – Conclusions and Future Work

This new design method is expected to provide value to the field in three ways. First, the new criteria were demonstrated to enable the *in silico* selection of DNA systems

with duplex-formation rates varying by 5% or less. As such this method enables the creation of devices with reproducibility and quality not previously possible. Second, the new algorithm is expected to both facilitate efficient generation of sequences according to our criteria, and influence the development of algorithms for future criteria. Third, the SeqEvo and DevPro software were built to be both widely applicable and usable by a typical advanced computer user. As such, it is expected that these tools will make the process of *in silico* sequence optimization available for anyone willing to learn to use them. This may be further amplified by the fact the tools are both publicly available for no cost and open source.

However, there remains substantial opportunity for additional work. For example, while it was demonstrated that the elimination of small intramolecular B-DNA type structures leads to kinetic uniformity, relatively little is known about how or why these structures lead to kinetic deviation. A systematic study of model systems with strategically introduced structures may lead to a better understanding of this relationship, and consequently the creation of improved fitness-criteria. As another example, it may be possible to greatly boost the efficiency of sequence generation by introducing new types of optimization algorithms. In principle, this would enable the creation of even larger fit systems.

APPENDIX A

Experimental Measurements

Section A.1 Generated Sequences

Table A.1 New sequences for the model system presented in Figure 2-5. The nomenculture for strand names is consistent with the disseration text.

Strand Name	Sequence (5'-3')
System TFS-1	
TFS-1_Strand-1	/5Cy3/TCC AAT CGC CCG TCG TAG GTG TGT CAG TAA TAA AGC AGT TCT CTC CAT G
TFS-1_Strand-2	CAT GGA GAG AAC TGC TTT ATT ACT GAC ACA CCT ACG ACG GGC GAT TGG A/3BHQ_1/
TFS-1_Strand-3	/5Cy5/TCC AAT CGC CCG TCG TAG GTG TGT CAG TAA TAA AGC AGT TC
System TFS-2	
TFS-2_Strand-1	/5Cy3/TAG TGT ATC CAA AGC CCG TAA GTC GCA GGT TCG TGT CAA TCT CTC CAT G
TFS-2_Strand-2	CAT GGA GAG ATT GAC ACG AAC CTG CGA CTT ACG GGC TTT GGA TAC ACT A/3BHQ_1/
TFS-2_Strand-3	/5Cy5/TAG TGT ATC CAA AGC CCG TAA GTC GCA GGT TCG TGT CAA TC
System TFS-3	
TFS-3_Strand-1	/5Cy3/TCG TAG TGT GTC AGC AAA GTC CAA TAG GTT CGC CCG TAA TCT CTC CAT G
TFS-3_Strand-2	CAT GGA GAG ATT ACG GGC GAA CCT ATT GGA CTT TGC TGA CAC ACT ACG A/3BHQ_1/
TFS-3_Strand-3	/5Cy5/TCG TAG TGT GTC AGC AAA GTC CAA TAG GTT CGC CCG TAA TC
System NFS-1	
NFS-1_Strand-1	/5Cy3/TTA TCG TCA CAG TTC GGT TCC AAA GGG CAA TCA GCG TAG TCT CTC CAT G
NFS-1_Strand-2	CAT GGA GAG ACT ACG CTG ATT GCC CTT TGG AAC CGA ACT GTG ACG ATA A/3BHQ_1/
NFS-1_Strand-3	/5Cy5/TTA TCG TCA CAG TTC GGT TCC AAA GGG CAA TCA GCG TAG TC
System NFS-2	
NFS-2_Strand-1	/5Cy3/TCG GCG TAA GCA ATA GGT TTC ACA ATC CCA GGT AGT CGT TCT CTC CAT G
NFS-2_Strand-2	CAT GGA GAG AAC GAC TAC CTG GGA TTG TGA AAC CTA TTG CTT ACG CCG A/3BHQ_1/
NFS-2_Strand-3	/5Cy5/TCG GCG TAA GCA ATA GGT TTC ACA ATC CCA GGT AGT CGT TC
System NFS-3	
NFS-3_Strand-1	/5Cy3/TGT AAA TCC CGT GCT AAA GTA TCG TCG CCA AGG TTC AGG TCT CTC CAT G
NFS-3_Strand-2	CAT GGA GAG ACC TGA ACC TTG GCG ACG ATA CTT TAG CAC GGG ATT TAC A/3BHQ_1/
NFS-3_Strand-3	/5Cy5/TGT AAA TCC CGT GCT AAA GTA TCG TCG CCA AGG TTC AGG TC
System SFS-1	
SFS-1_Strand-1	/5Cy3/TAA AAG TGT GTA AAA AAG TCC CGT GTC CGT GTG TCC GTC CCT CTC CAT G
SFS-1_Strand-2	CAT GGA GAG GGA CGG ACA CAC GGA CAC GGG ACT TTT TTA CAC ACT TTT A/3BHQ_1/
SFS-1_Strand-3	/5Cy5/TAA AAG TGT GTA AAA AAG TCC CGT GTC CGT GTG TCC GTC CC
System SFS-2	
SFS-2_Strand-1	/5Cy3/TCG TGT GTG TGT CCC GTA AAA GTA AAA AAG TCC CGT GTC CCT CTC CAT G
SFS-2_Strand-2	CAT GGA GAG GGA CAC GGG ACT TTT TTA CTT TTA CGG GAC ACA CAC ACG A/3BHQ_1/
SFS-2_Strand-3	/5Cy5/TCG TGT GTG TGT CCC GTA AAA GTA AAA AAG TCC CGT GTC CC

Strand Name	Sequence (5’-3’)
System SFS-3	
SFS-3_Strand-1	/5Cy3/TGT GTA AAA GTG TCC CGT GTC GTA AAA AAG TCC CGT GTC CCT CTC CAT G
SFS-3_Strand-2	CAT GGA GAG GGA CAC GGG ACT TTT TTA CGA CAC GGG ACA CTT TTA CAC A/3BHQ_1/
SFS-3_Strand-3	/5Cy5/TGT GTA AAA GTG TCC CGT GTC GTA AAA AAG TCC CGT GTC CC
System RND-1	
RND-1_Strand-1	/5Cy3/GTG TCA ACA CCT CGC TAG AGA TGG TGC GCT AAA TTA CGC TTC TCC ATG
RND-1_Strand-2	CAT GGA GAA GCG TAA TTT AGC GCA CCA TCT CTA GCG AGG TGT TGA CAC /3BHQ_1/
RND-1_Strand-3	/5Cy5/GTG TCA ACA CCT CGC TAG AGA TGG TGC GCT AAA TTA CGC T
System RND-2	
RND-2_Strand-1	/5Cy3/GAT TAG TCA TTA AGG GAT CGA CAC CAC GGG CTT CTT CCG ATC TCC ATG
RND-2_Strand-2	CAT GGA GAT CGG AAG AAG CCC GTG GTG TCG ATC CCT TAA TGA CTA ATC /3BHQ_1/
RND-2_Strand-3	/5Cy5/GAT TAG TCA TTA AGG GAT CGA CAC CAC GGG CTT CTT CCG A
System RND-3	
RND-3_Strand-1	/5Cy3/TCC TAT GTA CAG TCG TAC GGA CTA TTG CGG AAC CCT GAG ATC TCC ATG
RND-3_Strand-2	CAT GGA GAT CTC AGG GTT CCG CAA TAG TCC GTA CGA CTG TAC ATA GGA /3BHQ_1/
RND-3_Strand-3	/5Cy5/TCC TAT GTA CAG TCG TAC GGA CTA TTG CGG AAC CCT GAG A

Section A.2 Fluorescence Measurements and Kinetic Modeling

Measurements were organized into experiments consisting of up to three samples. Sample 1 contained a dye only control (black circles, reactants: Strand-1, labeled “1”). Sample 2 contained the duplex formation (DF) reaction (red triangles, reactants: Strand-1 and Strand-2, labeled “1 & 2”). Sample 3 contained the strand displacement (SD) reaction (green squares, reactants: Strand-1 and Strand-2/Strand-3 complex, labeled “1 & 2:3”). Most data has three samples present, with the exception of the TFS-3 samples at 20 and 40 °C. These experiments were repeated in triplicate in order to study the reproducibility of the measurement process. As a result, experiments 32,33,34, 39, 40, and 41 contain measurements of the duplex-formation rates. Experiments 35, 36, 37, 42, 43, and 44 contain measurements of the strand-displacement rates.

Table A.2 System, temperature, and experiment number for the 82 sets of fluorescence measurements. Each experiment consisted of up to three samples, including a dye only control, duplex formation reaction, and strand displacement reaction. Experiments 32-37 and 39-44 were used to study the reproducibility of the measurement process, and contain only the duplex-formation reaction or the strand-displacement reaction as a consequence.

System	Temperature (°C)	Experiment No.
RND-1	10	1
	20	2
	30	3
	40	4
	50	5
	60	6
RND-2	10	7
	20	8
	30	9
	40	10
	50	11
	60	12
RND-3	10	13
	20	14
	30	15
	40	16
	50	17
	60	18
TFS-1	10	19
	20	20
	30	21
	40	22
	50	23
	60	24
TFS-2	10	25
	20	26
	30	27
	40	28
	50	29

System	Temperature (°C)	Experiment No.
	60	30
TFS-3	10	31
	20	32
		33
		34
		35
		36
		37
	30	38
	40	39
		40
		41
		42
		43
		44
	50	45
	60	46
SFS-1	10	47
	20	48
	30	19
	40	50
	50	51
	60	52
SFS-2	10	53
	20	54
	30	55
	40	56
	50	57
	60	58
SFS-3	10	59
	20	60
	30	61
	40	62
	50	63

System	Temperature (°C)	Experiment No.
	60	64
NFS-1	10	65
	20	66
	30	67
	40	68
	50	69
	60	70
NFS-2	10	71
	20	72
	30	73
	40	74
	50	75
	60	76
NFS-3	10	77
	20	78
	30	79
	40	80
	50	81
	60	82

Each two-page report contains the following graphs:

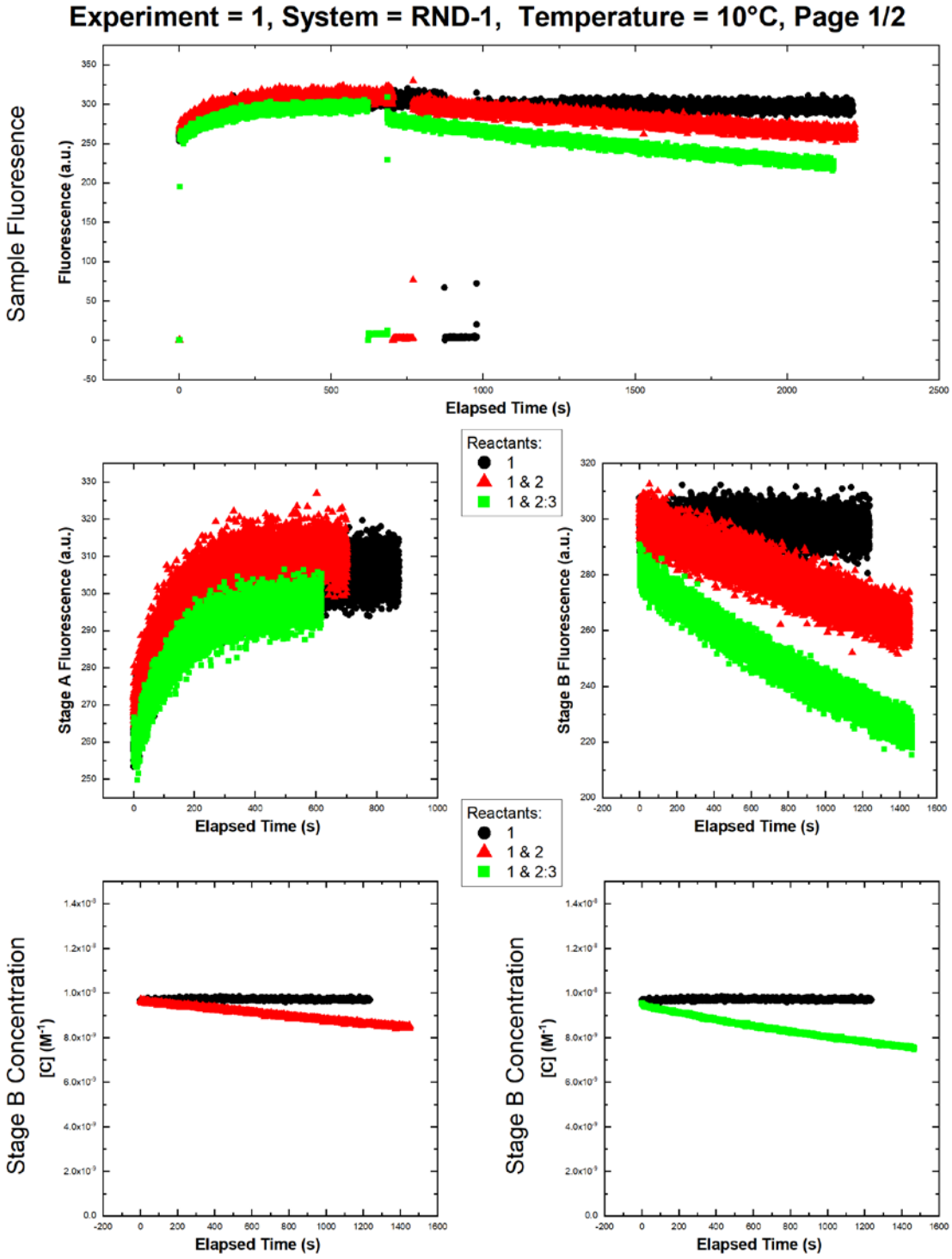
- 1st Page
    - A plot of the recorded fluorescence for each of the three samples. Each sample was approximately 1 mL of buffer/strand solution in a 1cm x 1cm x 4cm cuvette. Samples began with “Strand-1” slightly above 10nM concentration (time 0 in the red RND-1 / 10 °C / “1 & 2” sample below). Sample fluorescence was monitored as the sample came to the same temperature as the sample holder (time 0 to ~750s for the red RND-1 / 10 °C / “1 & 2” sample below). This was referred to as the first stage of the experiment.

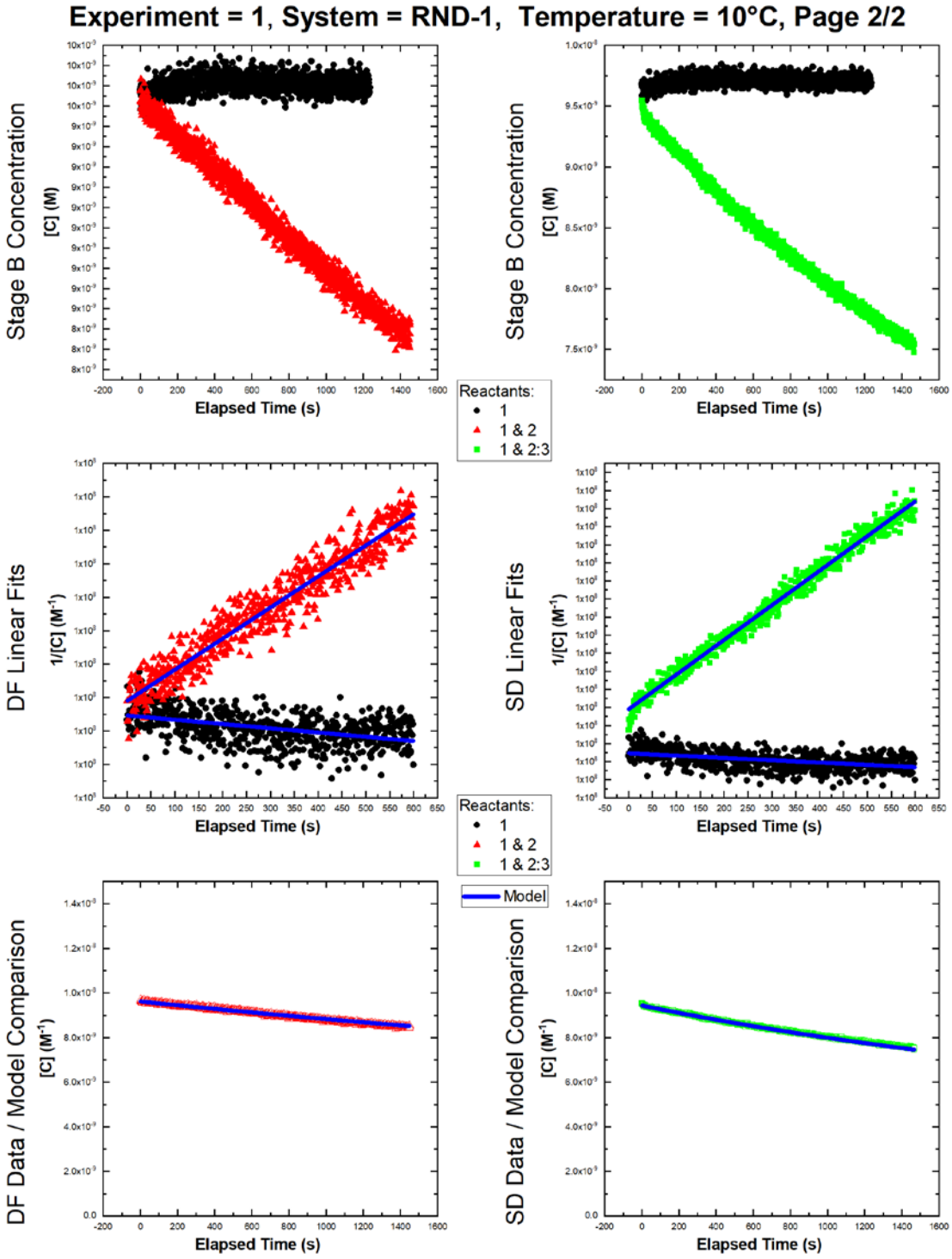
Samples were then removed from the machine, during which the fluorescence dropped to approximately zero. While samples were removed, the second reactant or an equivalent amount of buffer were added to the sample and the sample was mixed using a pipette. Following injection and mixing, reactant concentrations were 10nM. Samples were returned to the machine and fluorescence was monitored for a minimum of 10 minutes (time > 800s for the red RND-1 / 10 °C / "1 & 2" sample below). This was referred to as the second stage of the reaction. During this stage a decrease in fluorescence is observed as the fluorescent dye localizes with the quenching molecule. This is expected to occur for both reaction samples, but not the control sample (black circles). The stability of the control sample provides confidence that the spectrophotometer is working as expected and that there are no additional factors leading to fluorescence change during the experiment.

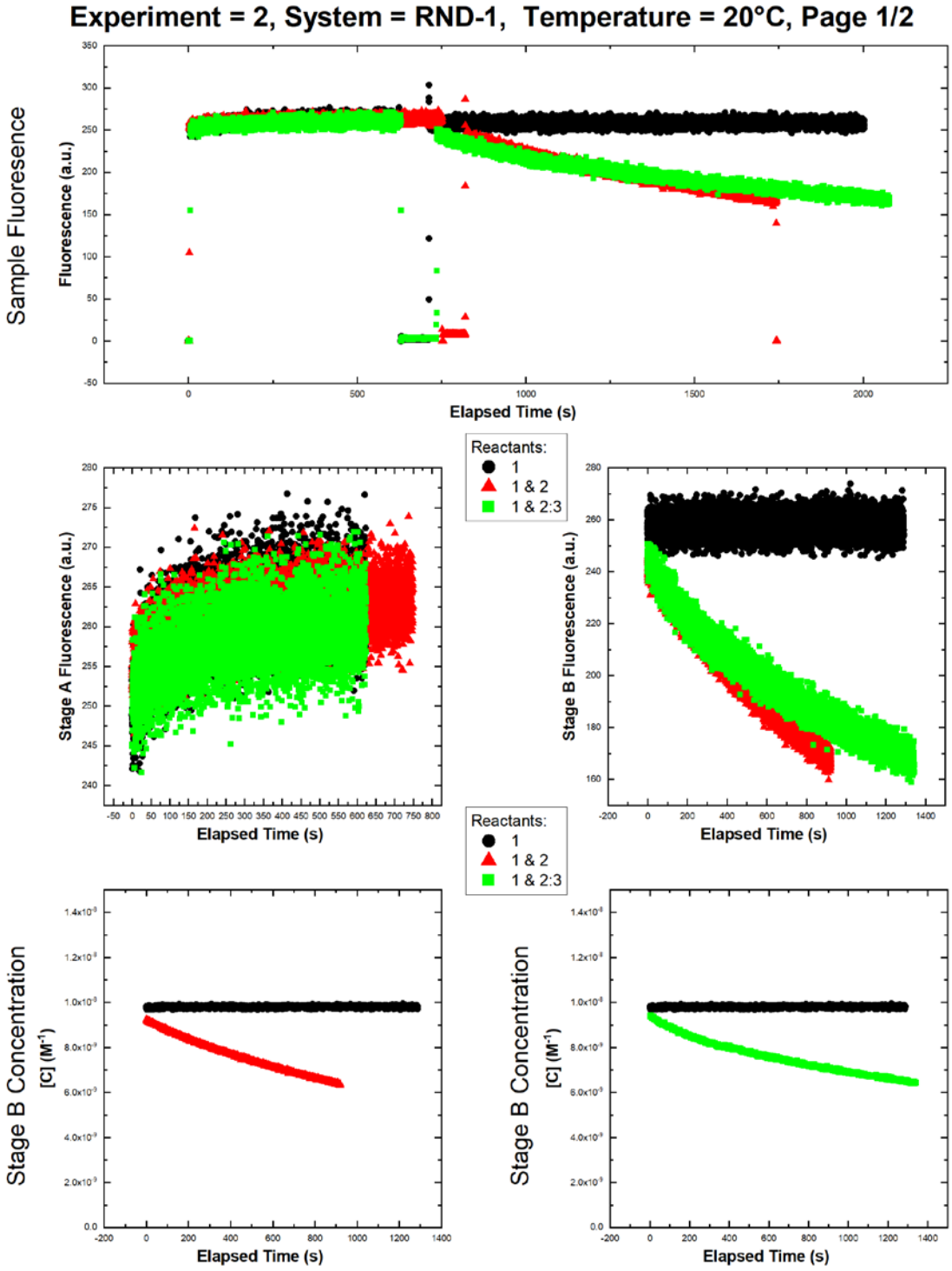
- Two plots of the fluorescent data extracted from stage 1 (sample temperature stabilization) and stage 2 (target reaction).
- A plot of reactant C's concentration vs time for the duplex-formation and control samples.
- A plot of reactant C's concentration vs time for the strand-displacement and control samples.

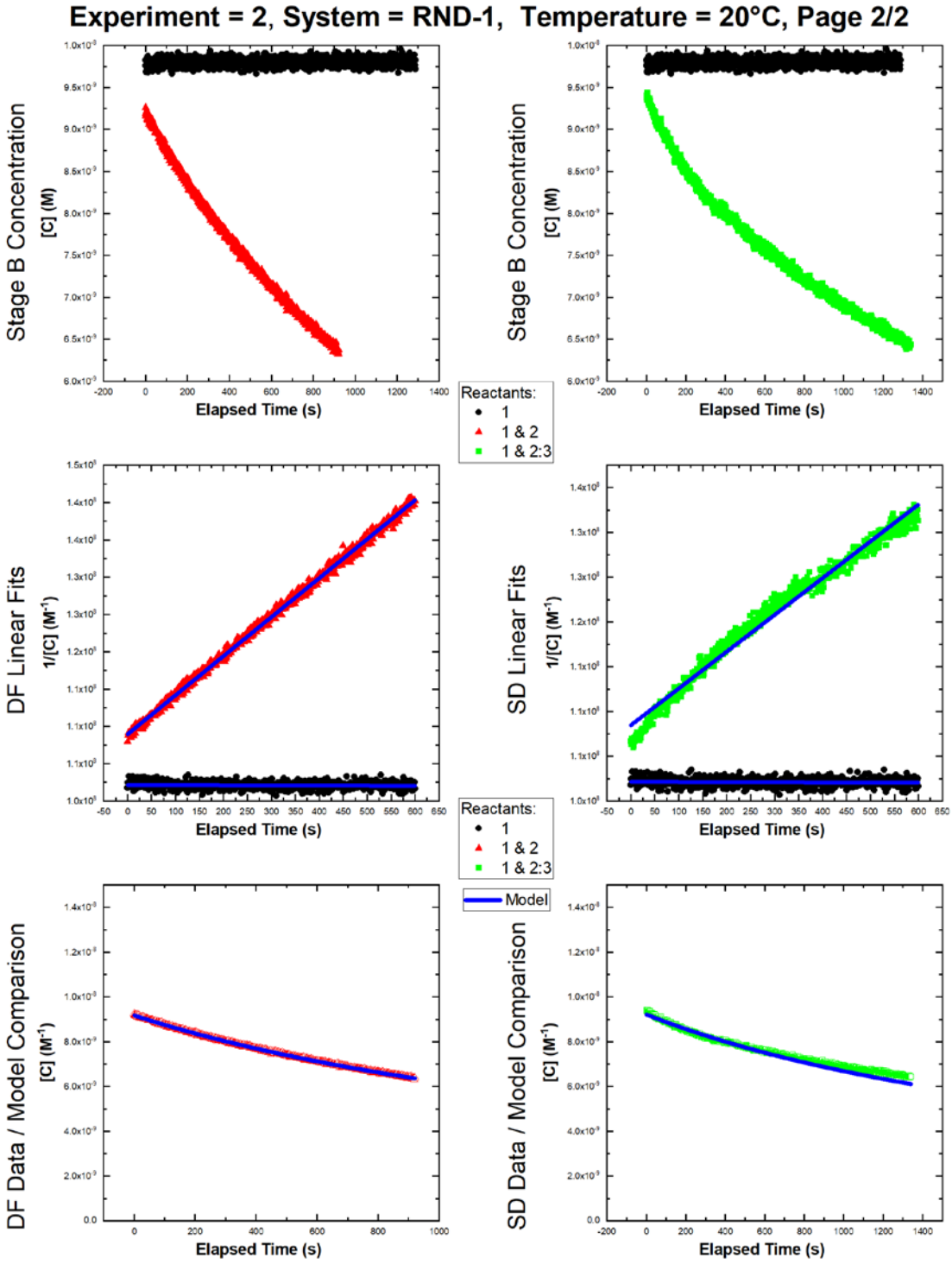
- 2nd Page
  - A duplication of the reactant concentrations vs time for both reactions and samples.

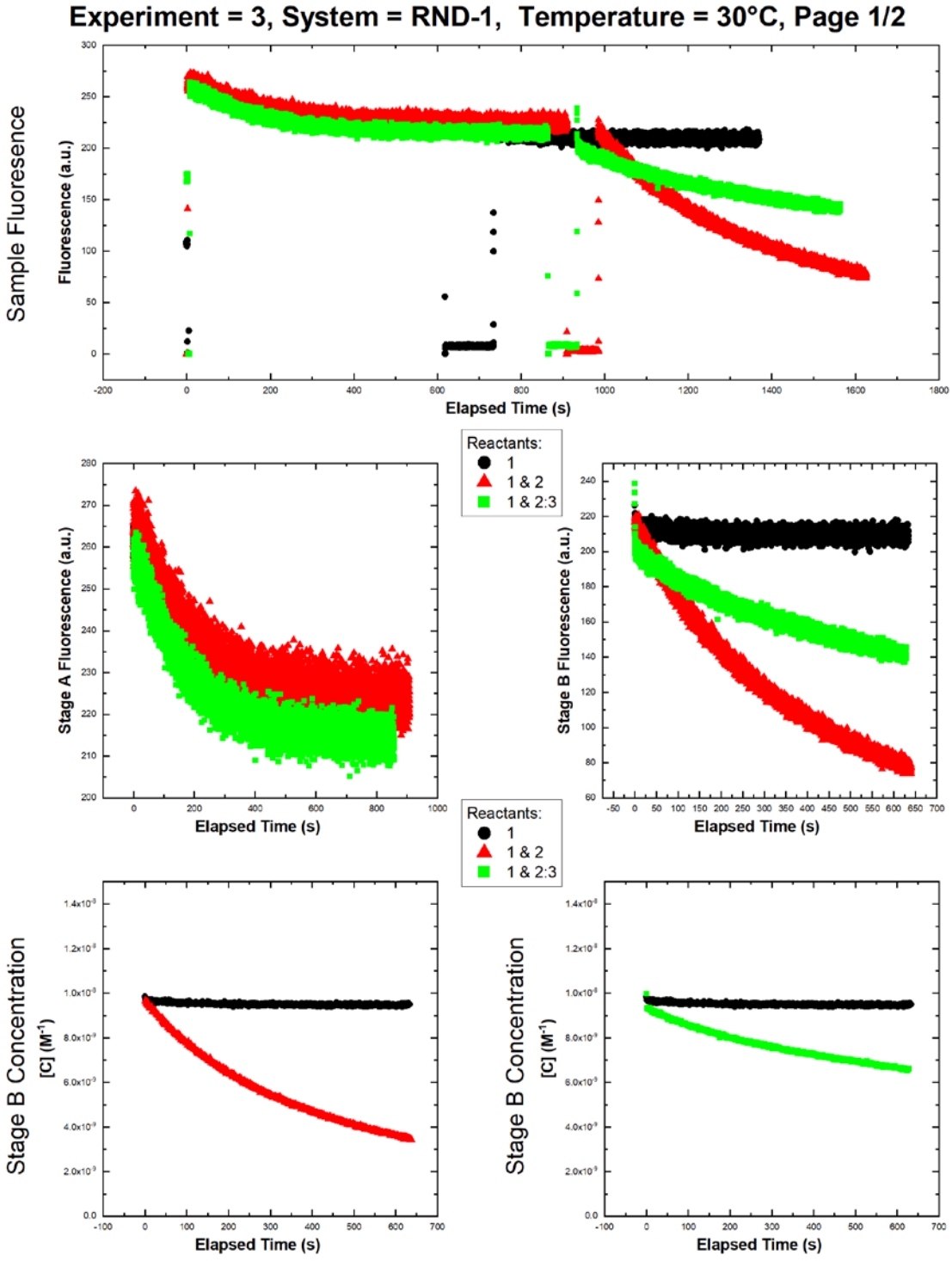
- Two plots of the inverse reactant concentrations from time 0 to reaction half-completion, or the first 600s for slower reactions. A linear fit was applied to and overlaid on this data (blue line). The slope of this linear fit is equal to the bimolecular reaction rate describing the reaction.
- A duplication of the reactant concentrations vs time plots with the model overlaid and extrapolated to predict the full range of data (blue trace).

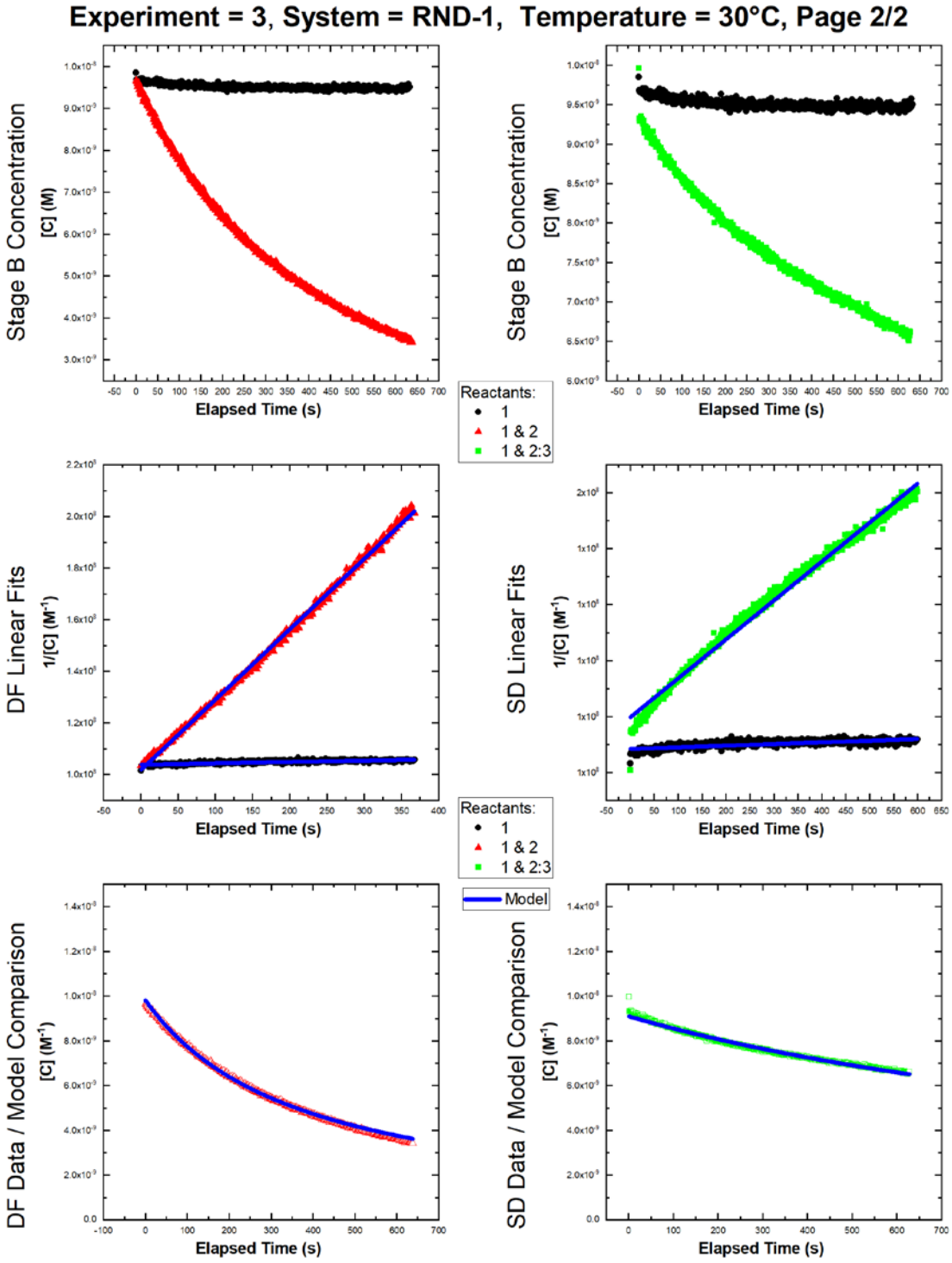

Experiment = 1, System = RND-1, Temperature = 10°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

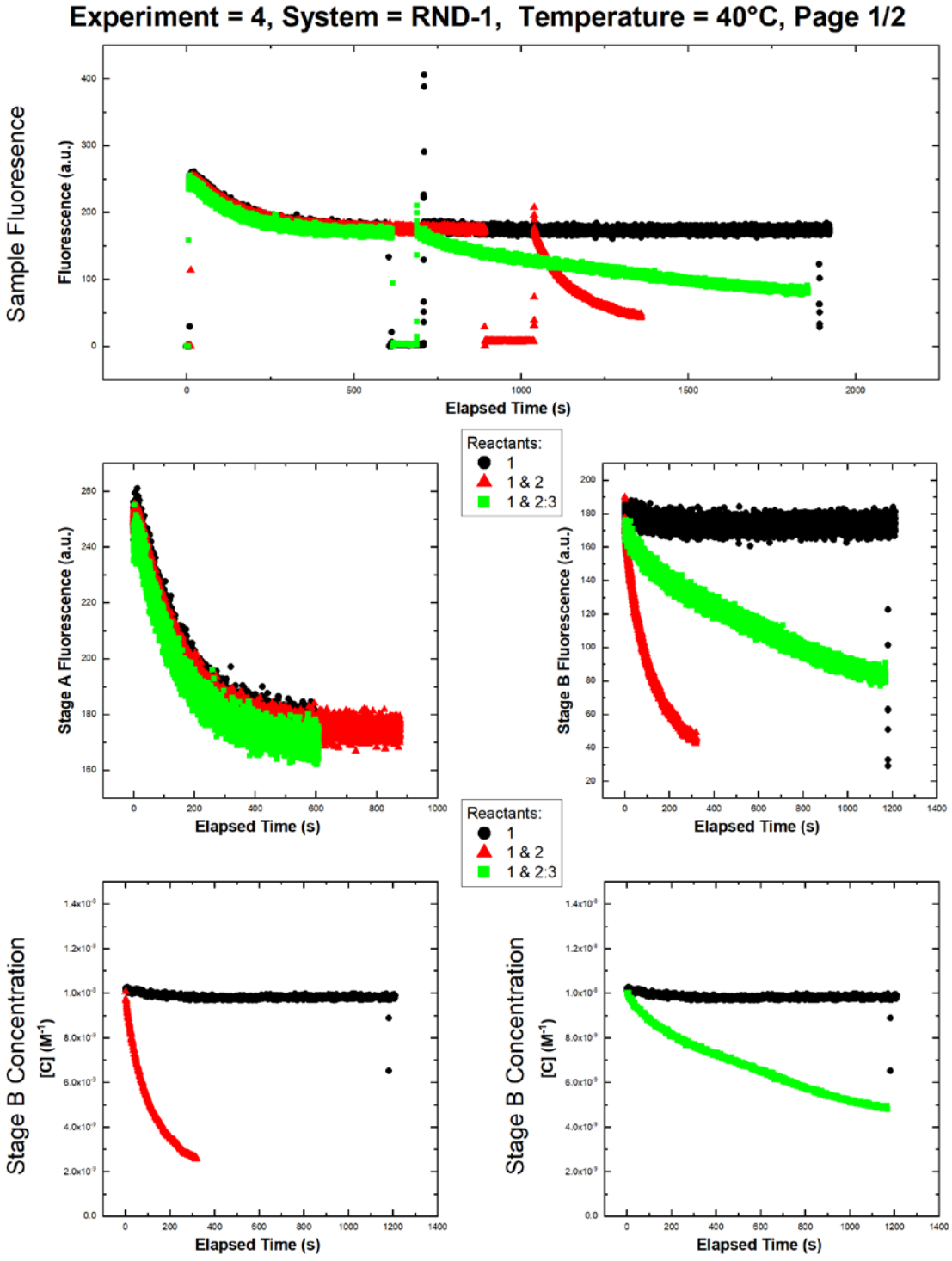

Experiment = 1, System = RND-1, Temperature = 10°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

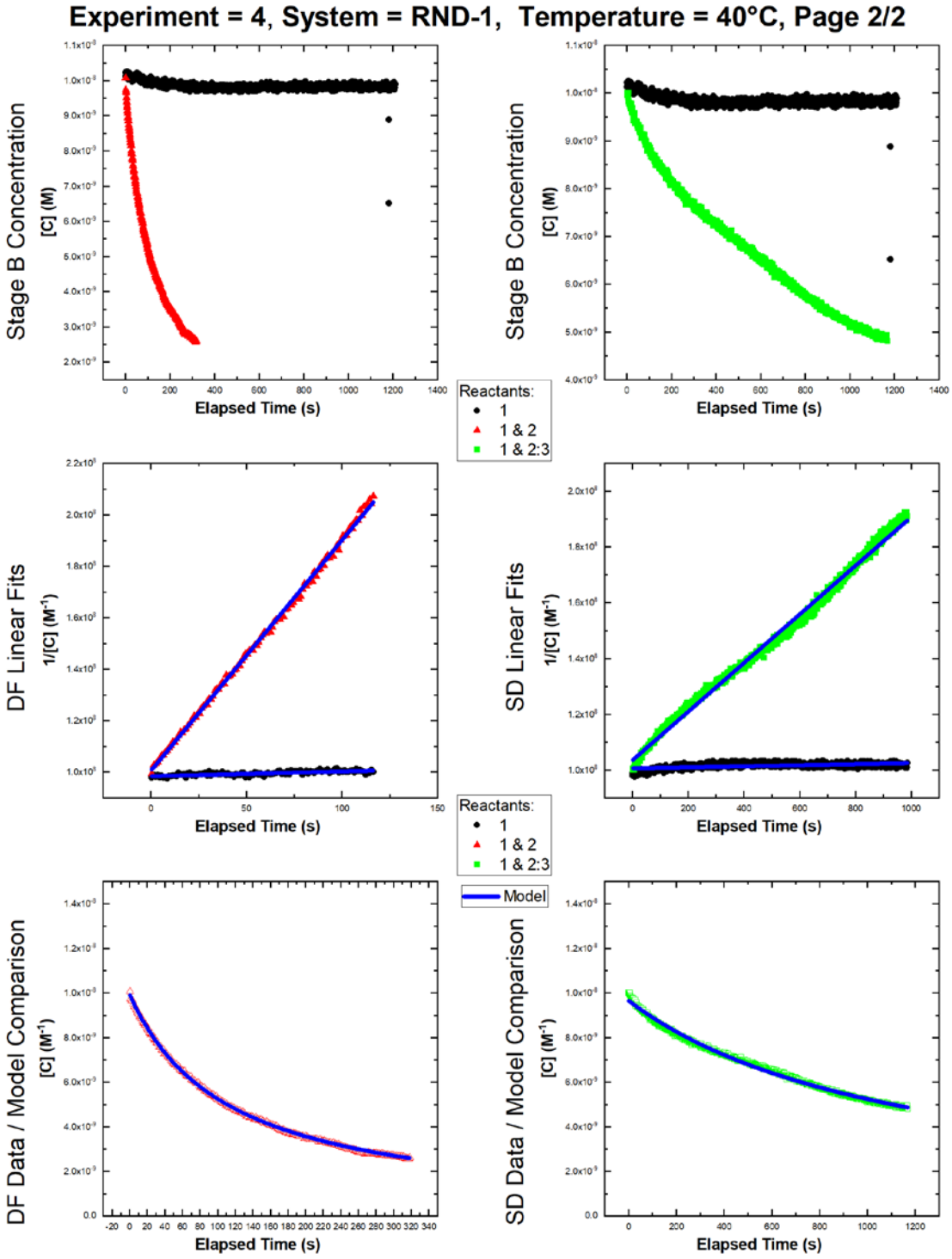

Experiment = 2, System = RND-1, Temperature = 20°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

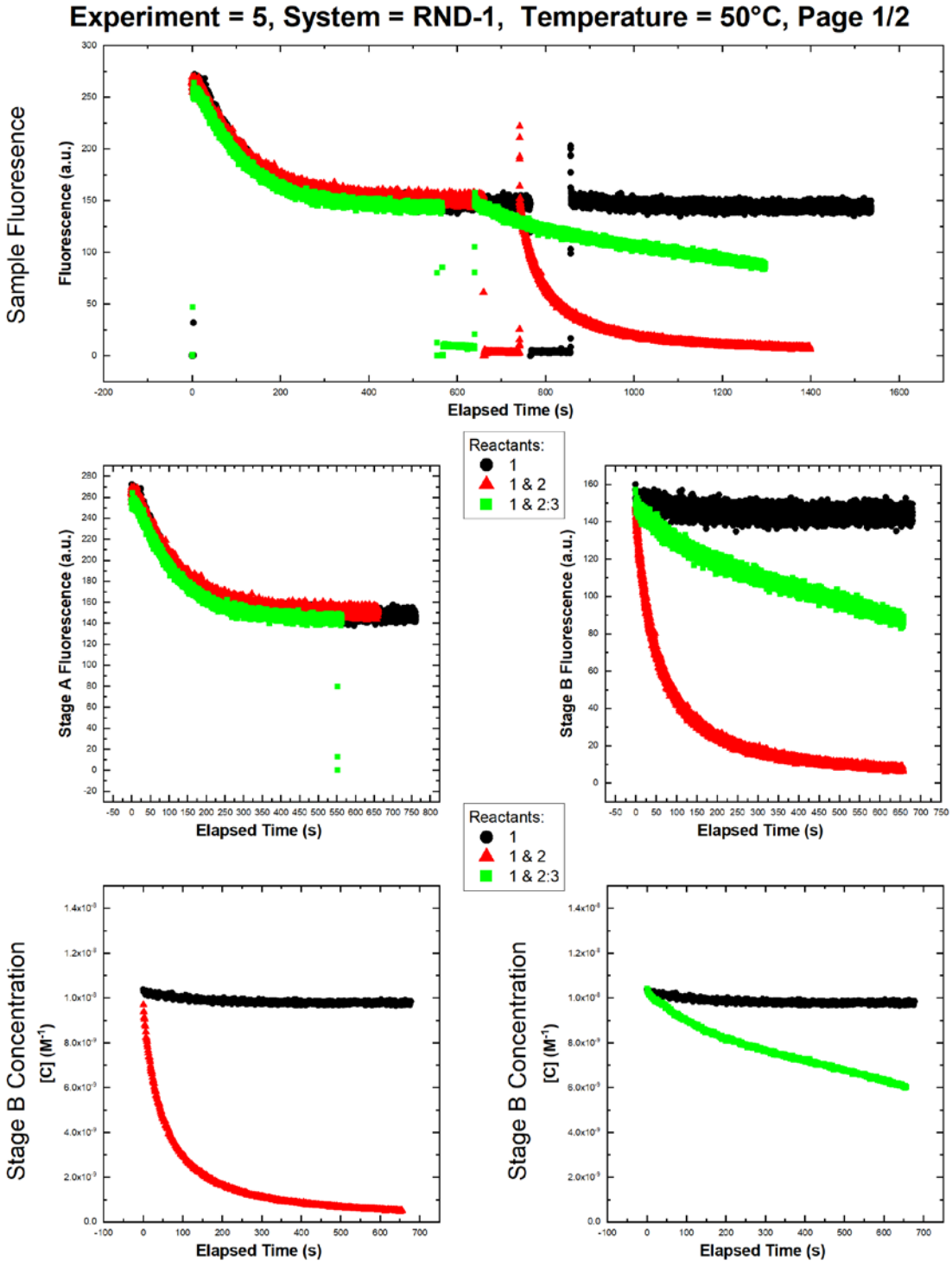

Experiment = 2, System = RND-1, Temperature = 20°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

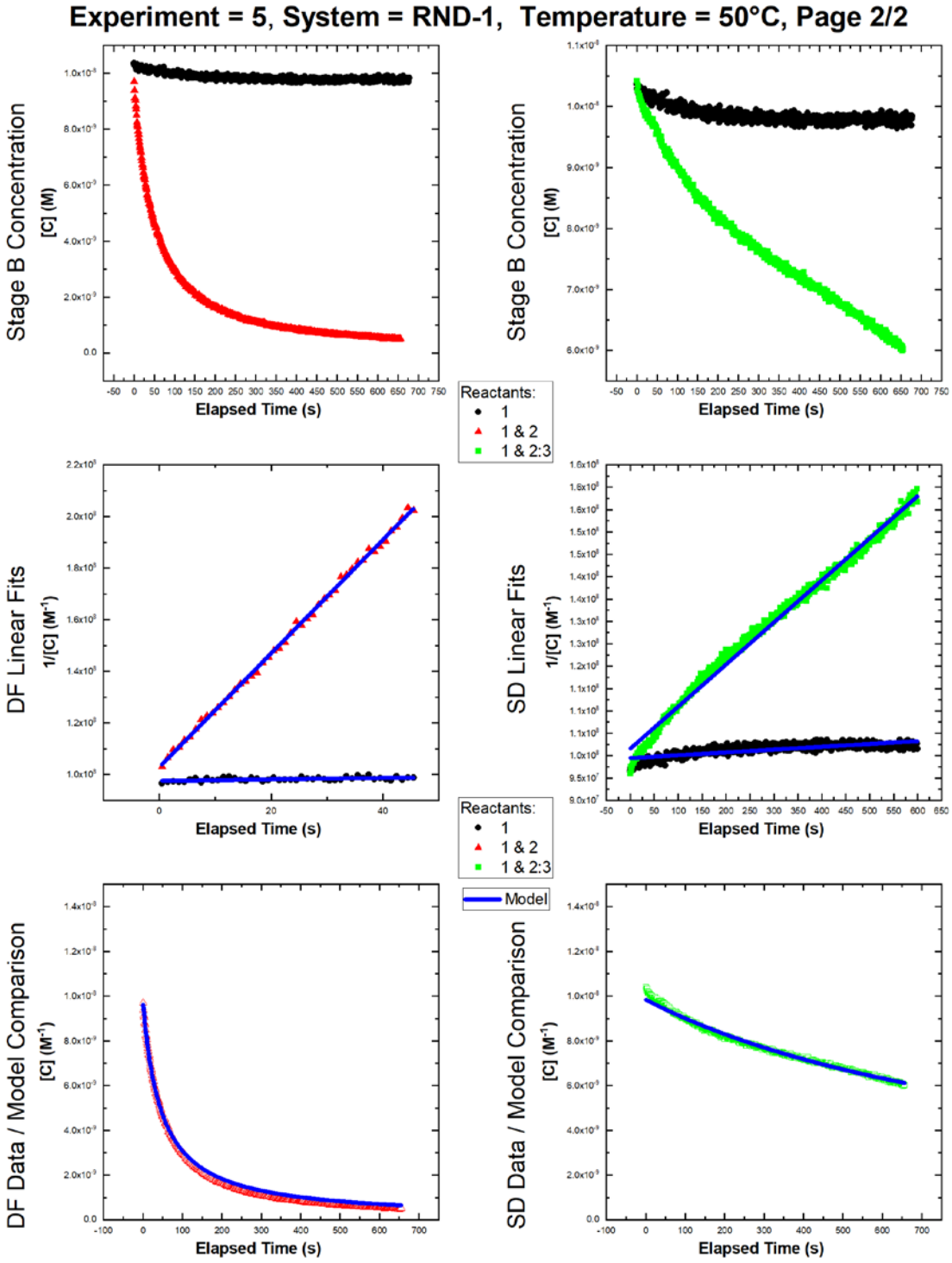

Experiment = 3, System = RND-1, Temperature = 30°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

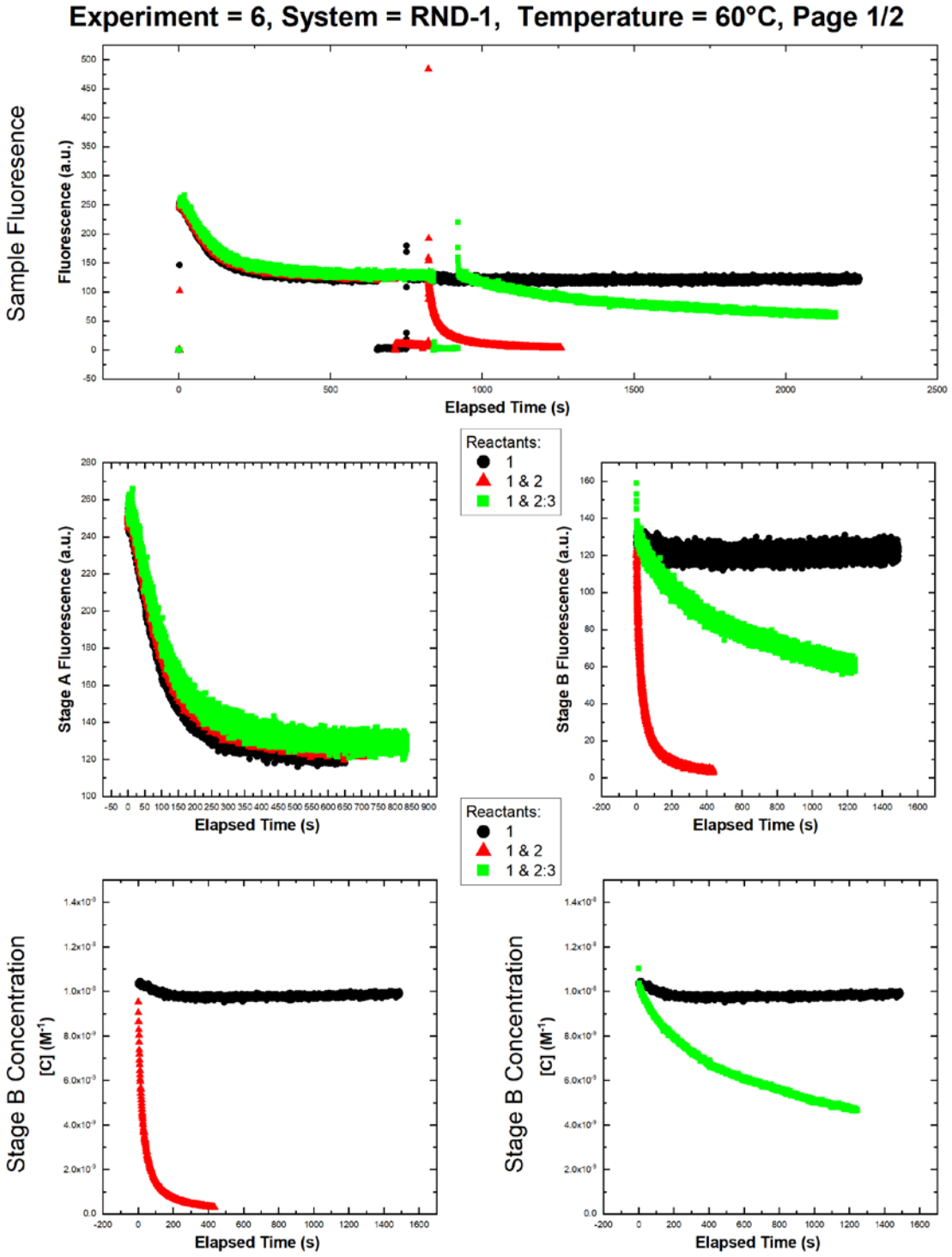

Experiment = 3, System = RND-1, Temperature = 30°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

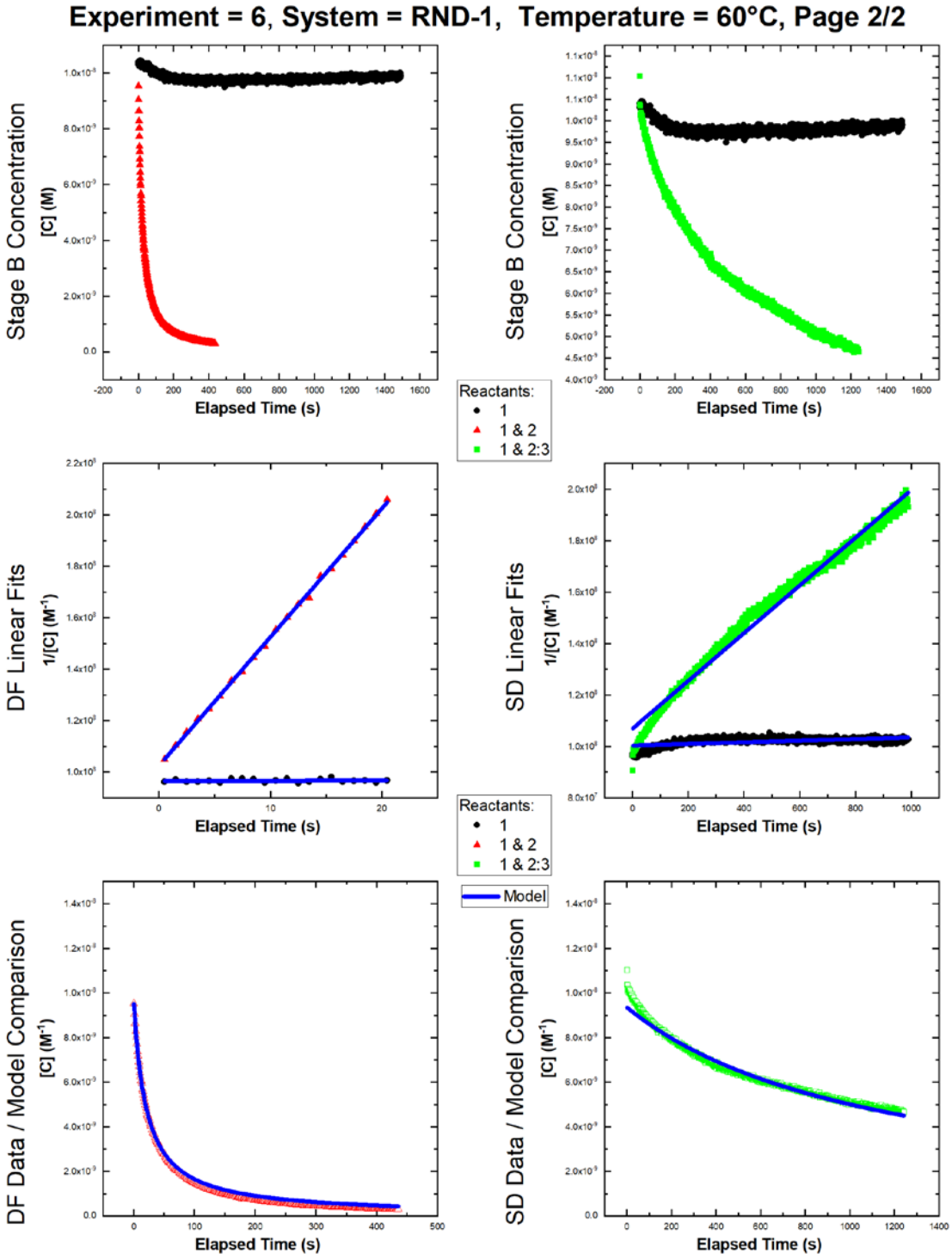

Experiment = 4, System = RND-1, Temperature = 40°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

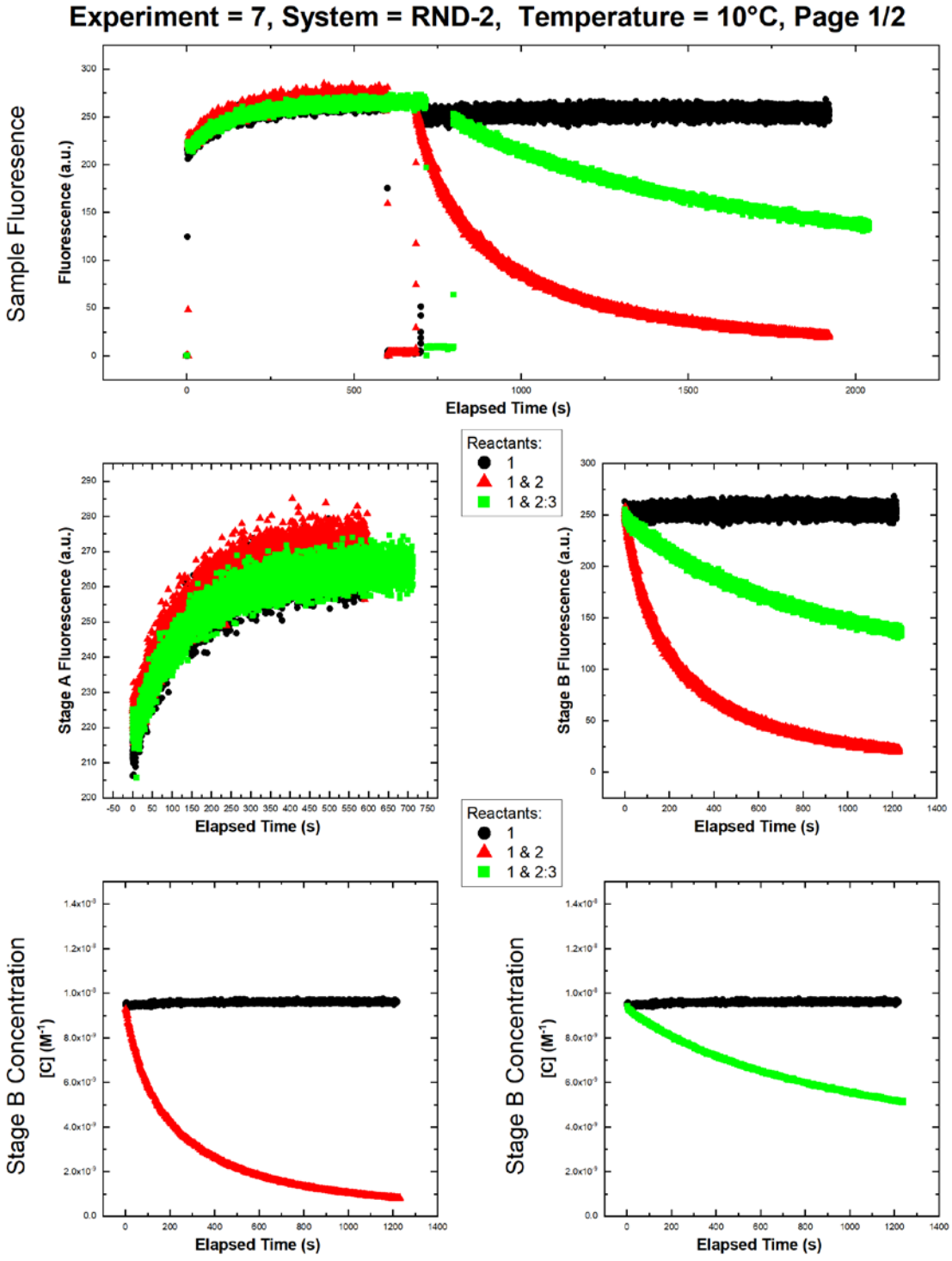

Experiment = 4, System = RND-1, Temperature = 40°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

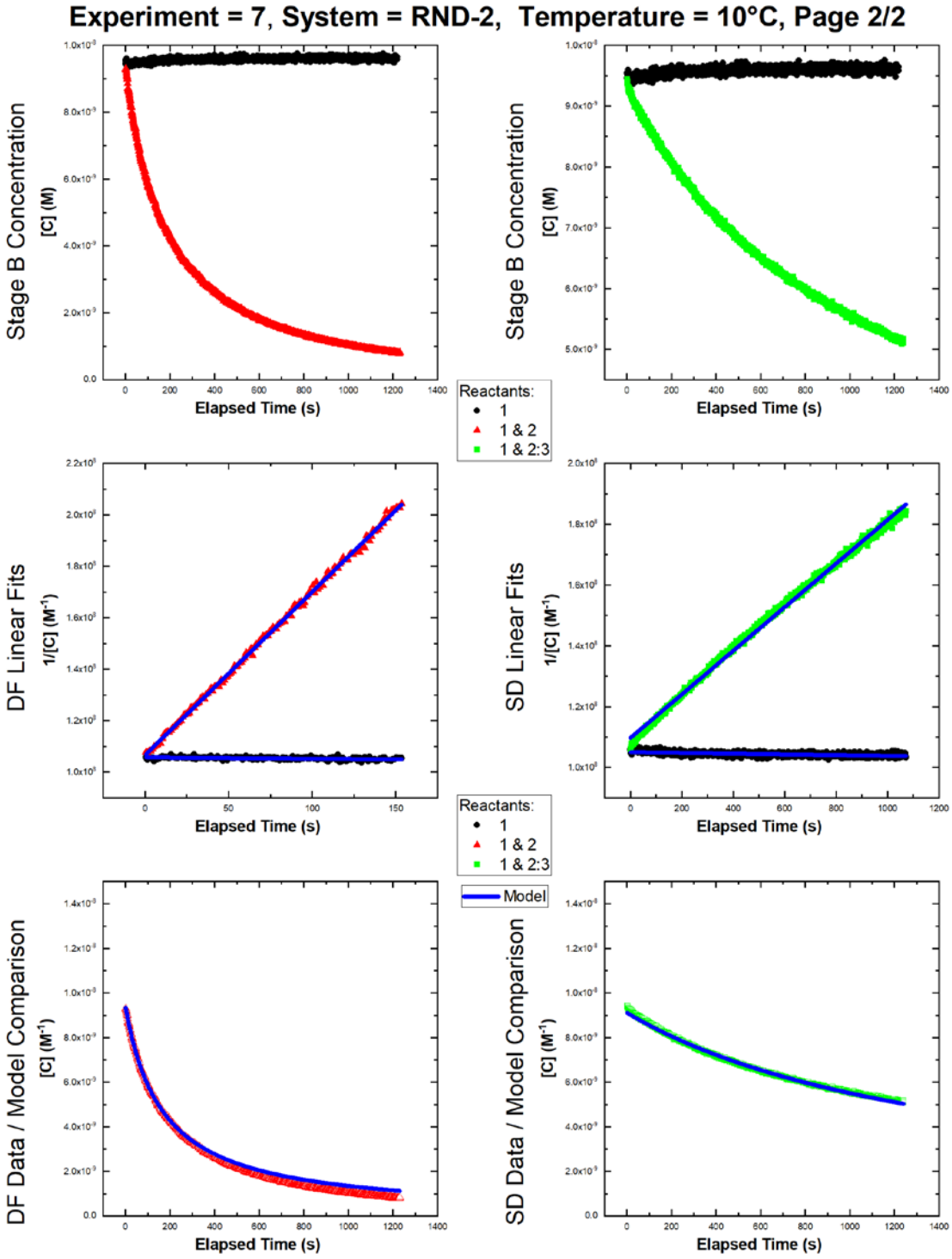

Experiment = 5, System = RND-1, Temperature = 50°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

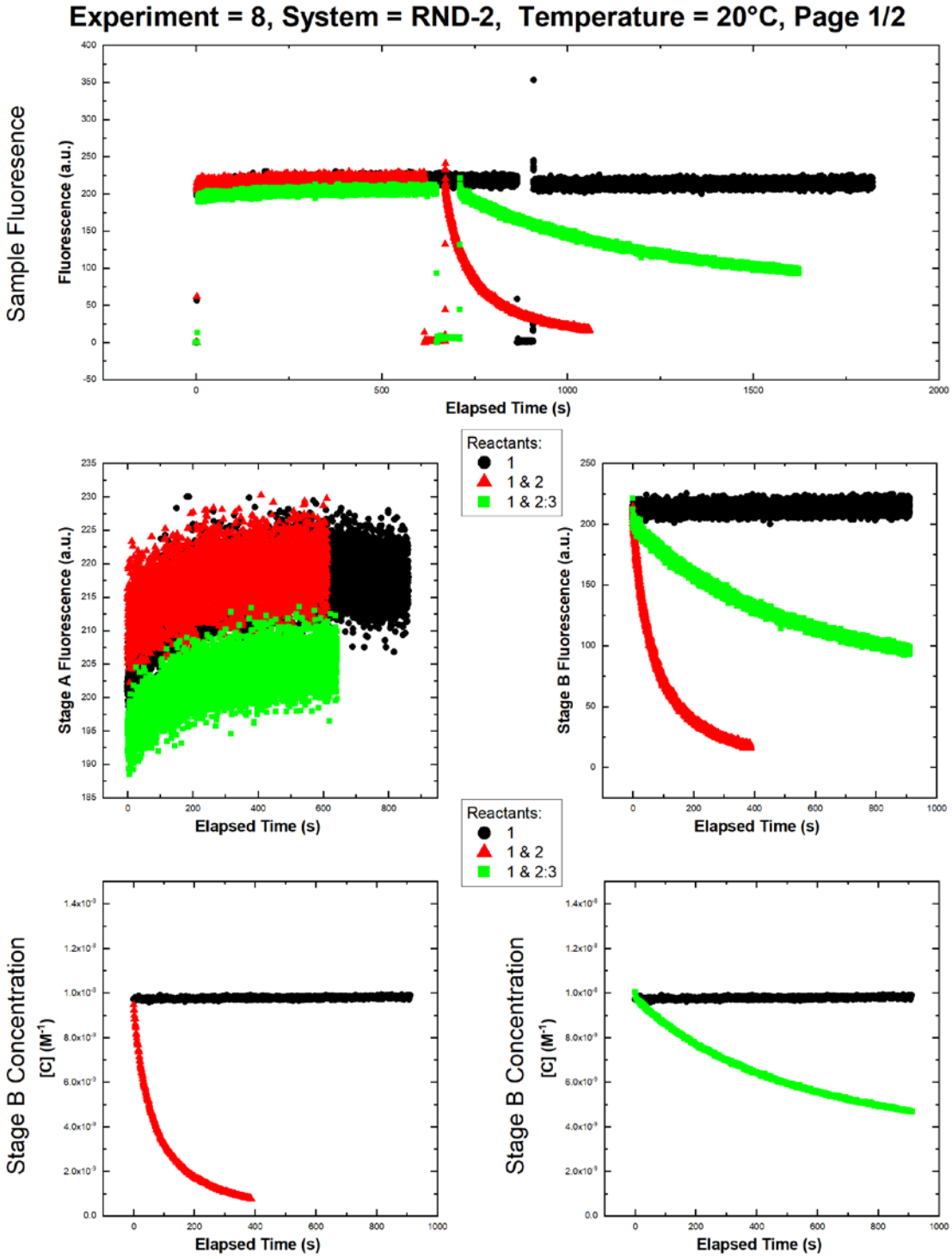

Experiment = 5, System = RND-1, Temperature = 50°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

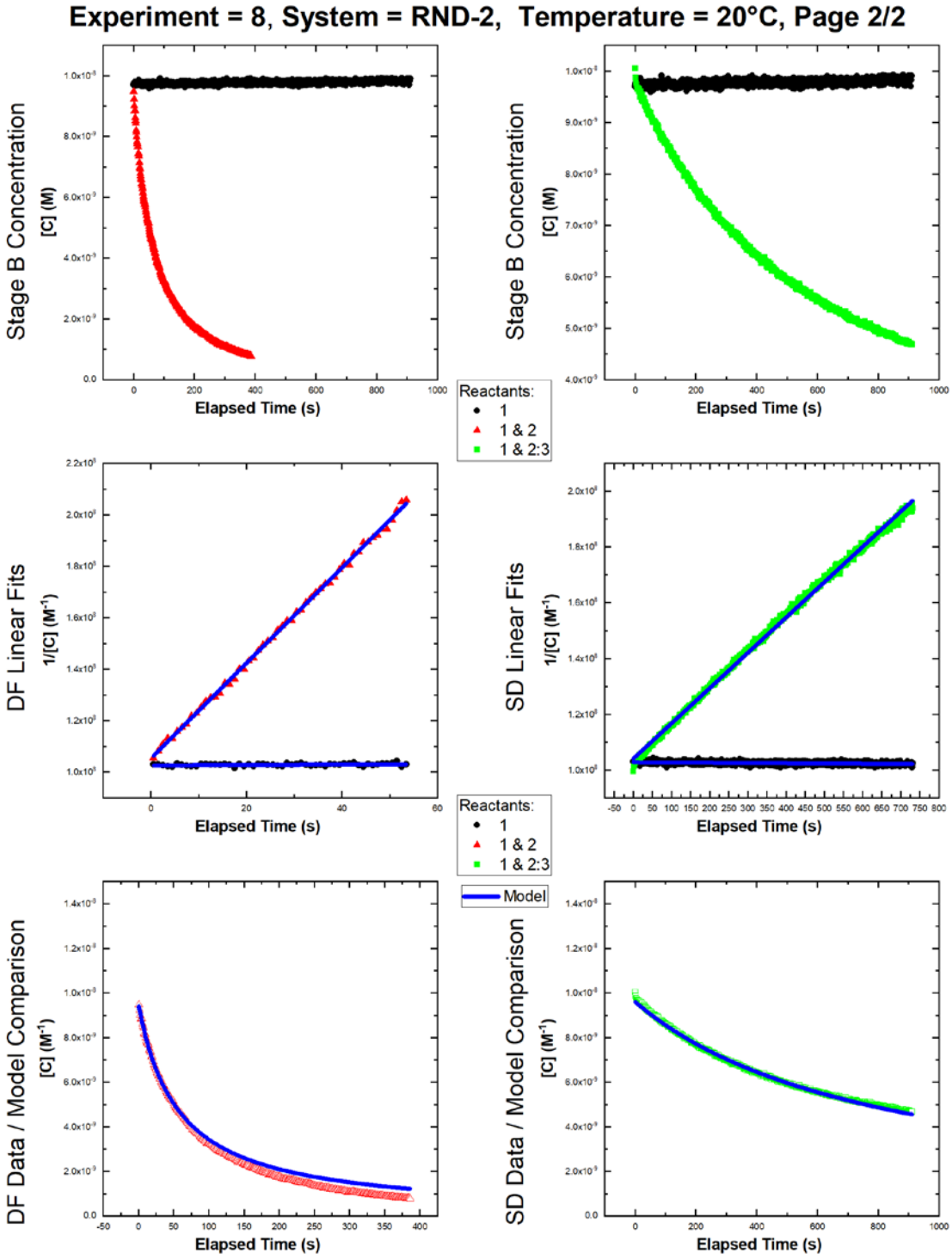

Experiment = 6, System = RND-1, Temperature = 60°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

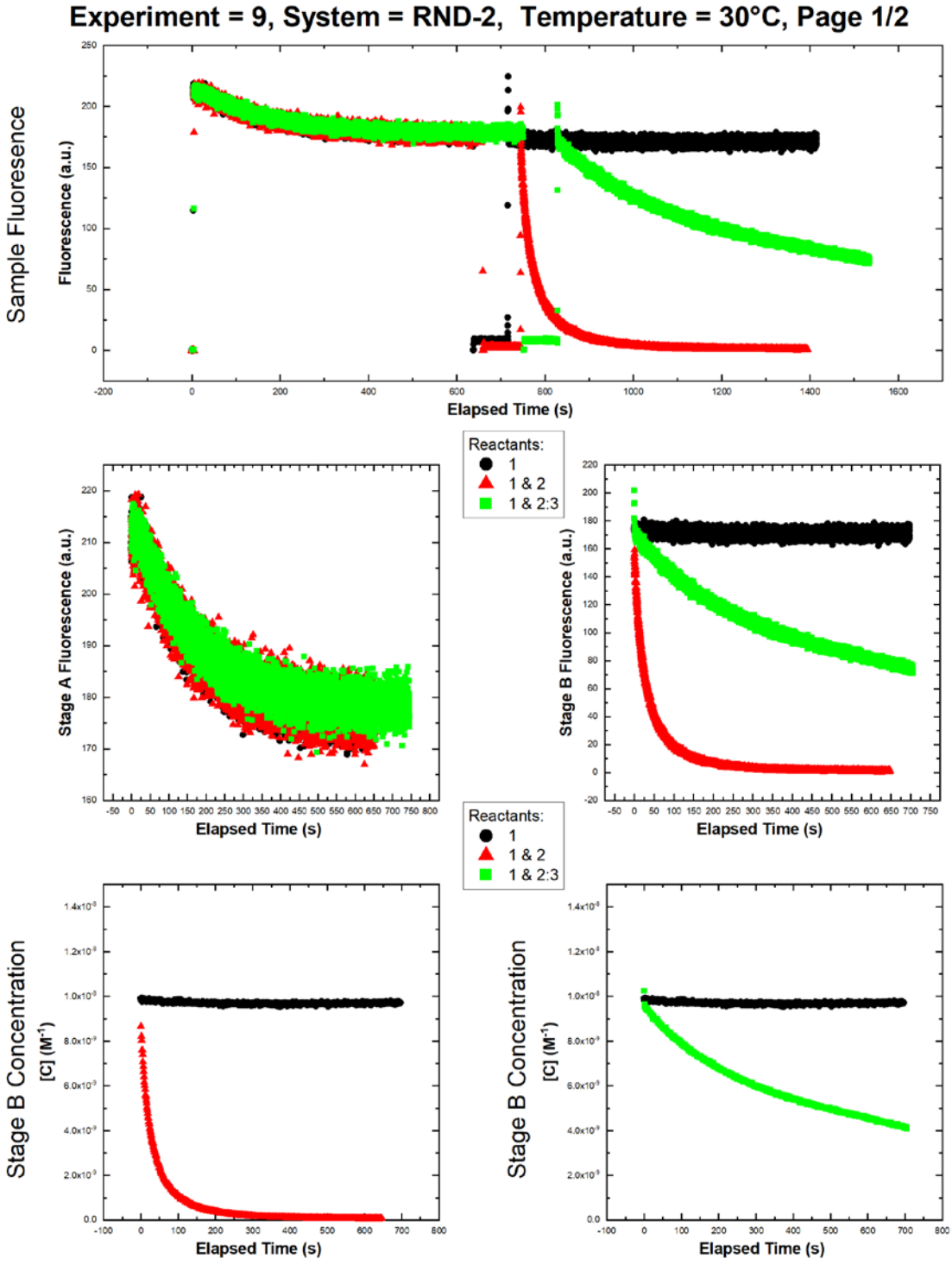

Experiment = 6, System = RND-1, Temperature = 60°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
[C] (M-1)
SD Data / Model Comparison

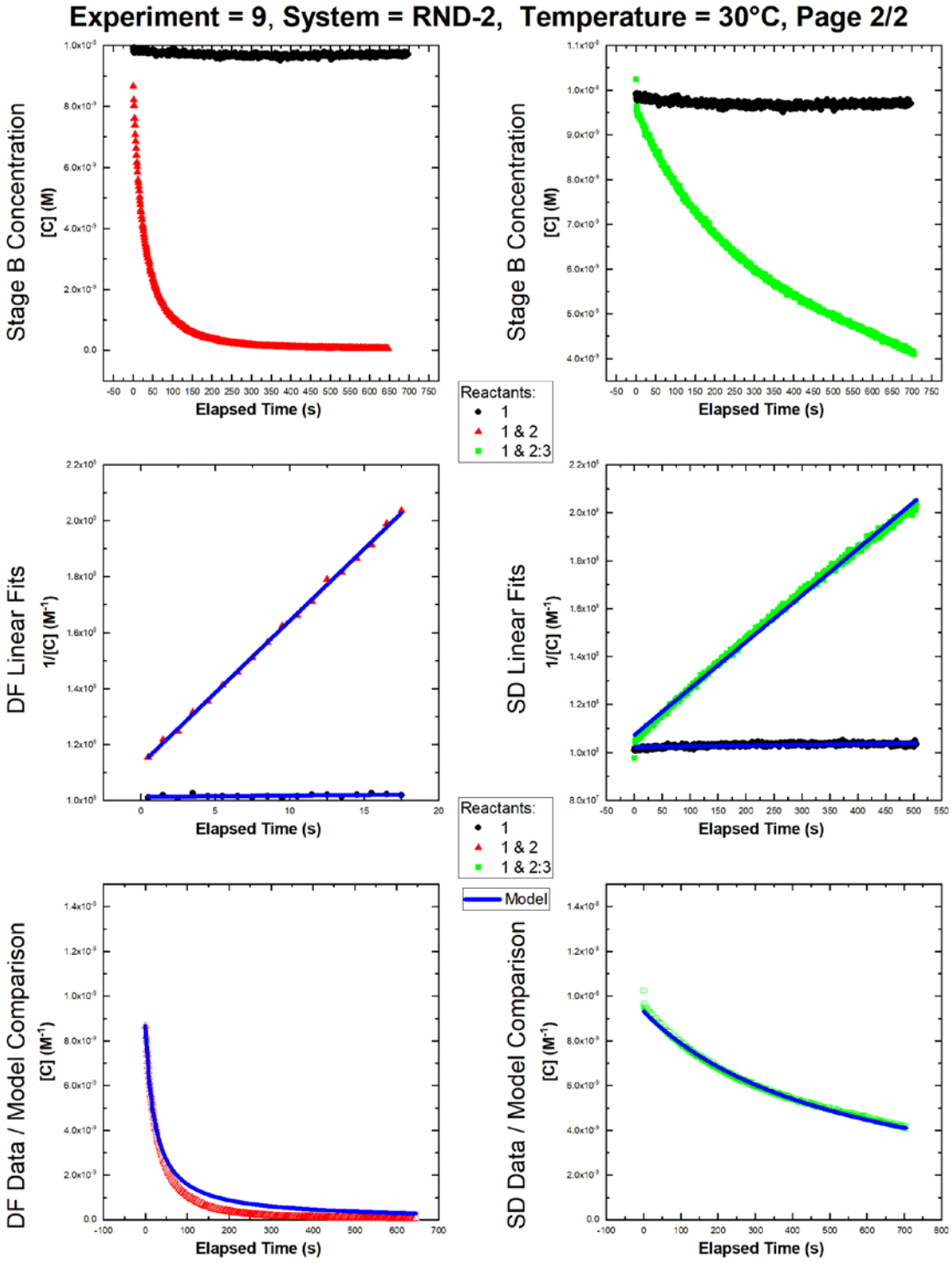

Experiment = 7, System = RND-2, Temperature = 10°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

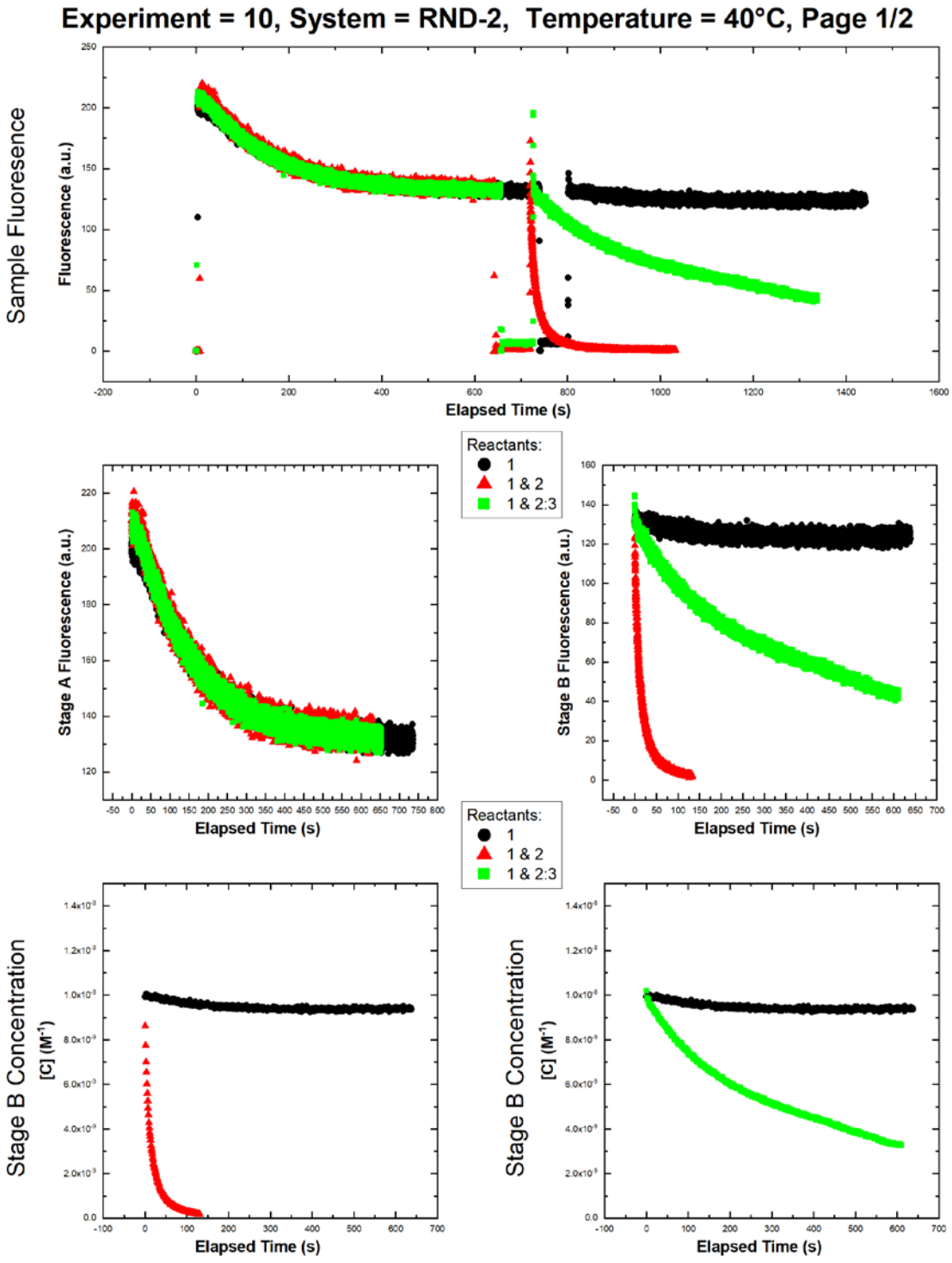

Experiment = 7, System = RND-2, Temperature = 10°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

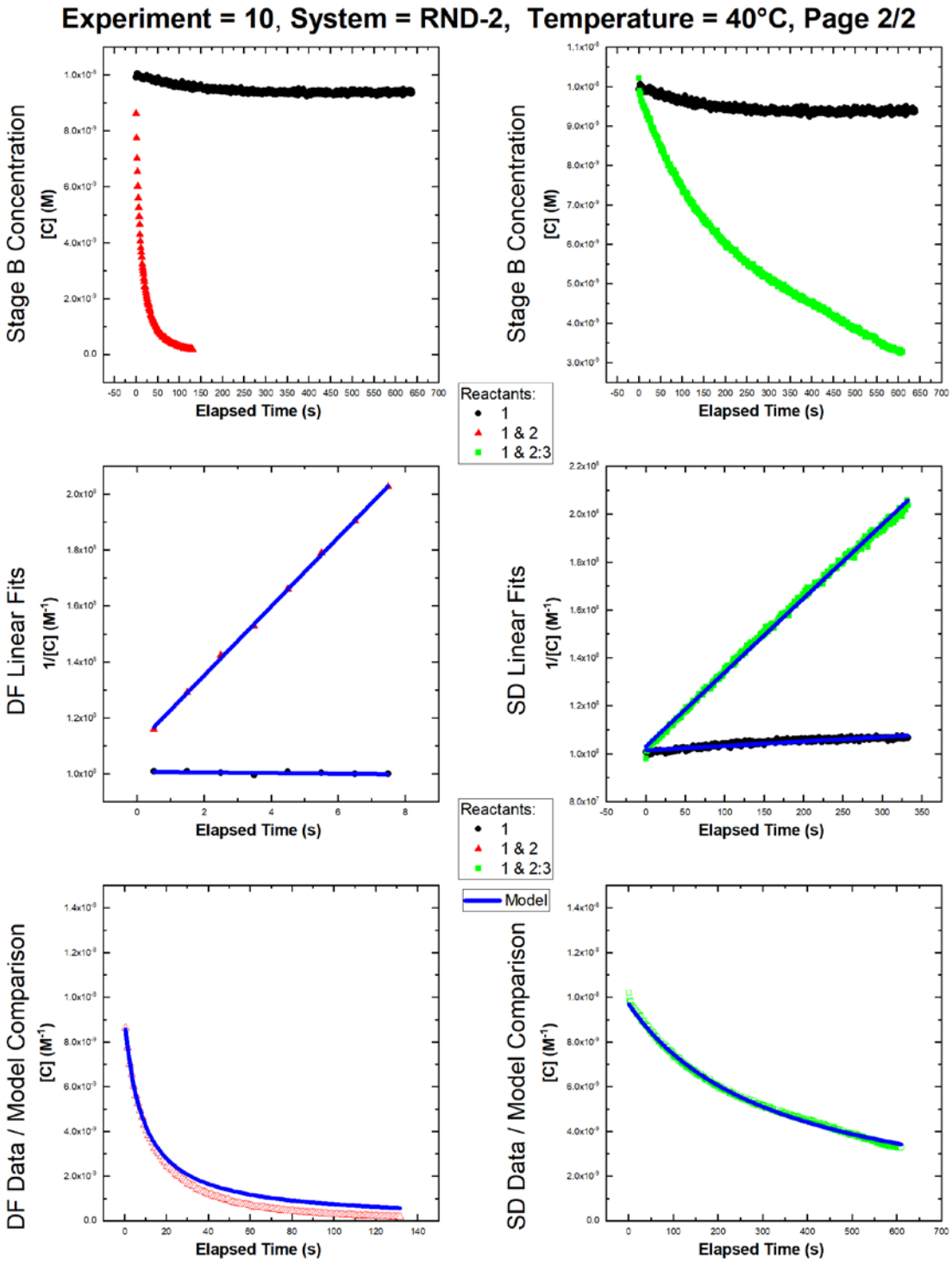

Experiment = 8, System = RND-2, Temperature = 20°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

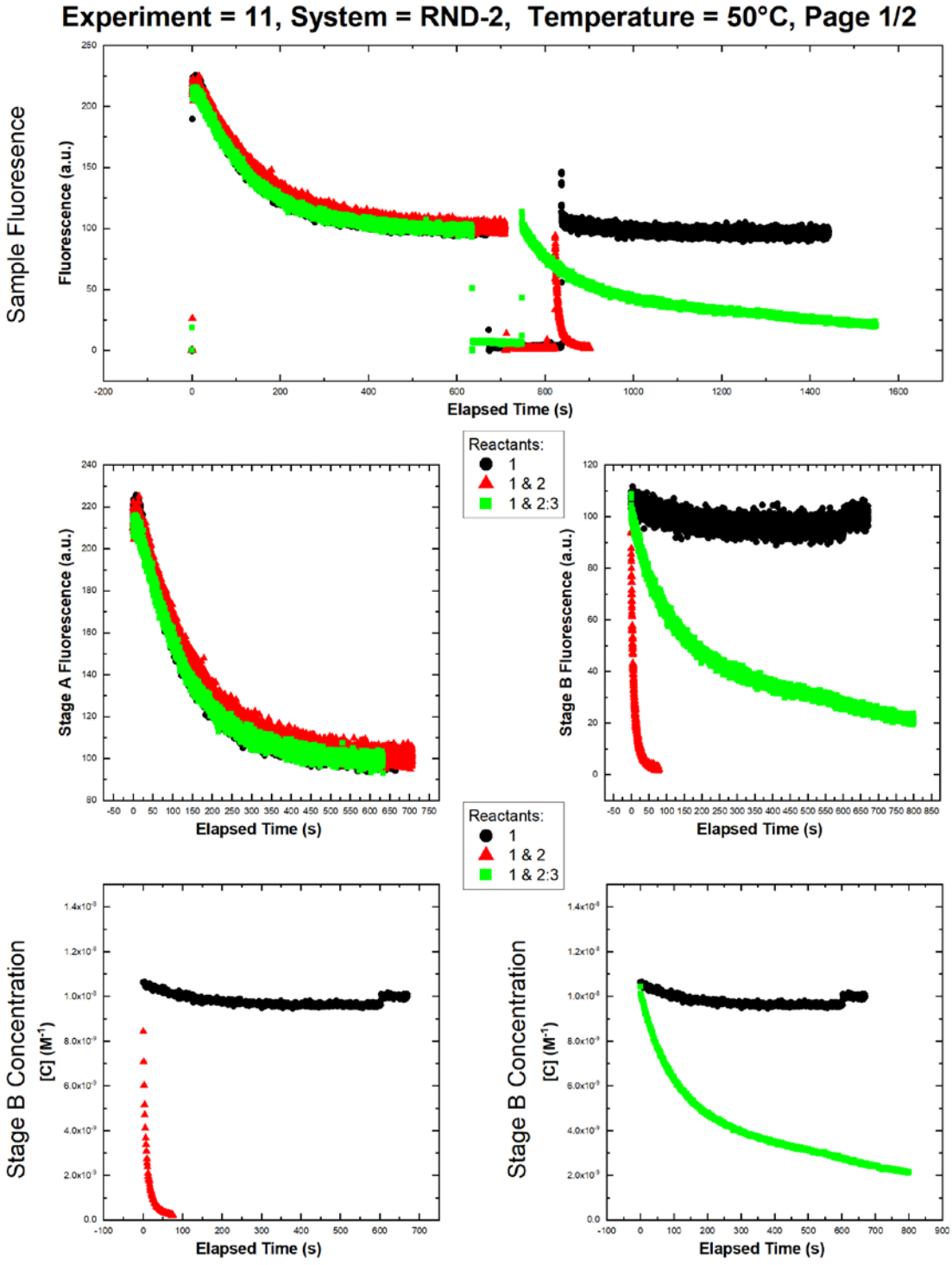

Experiment = 8, System = RND-2, Temperature = 20°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
[C] (M-1)
SD Data / Model Comparison

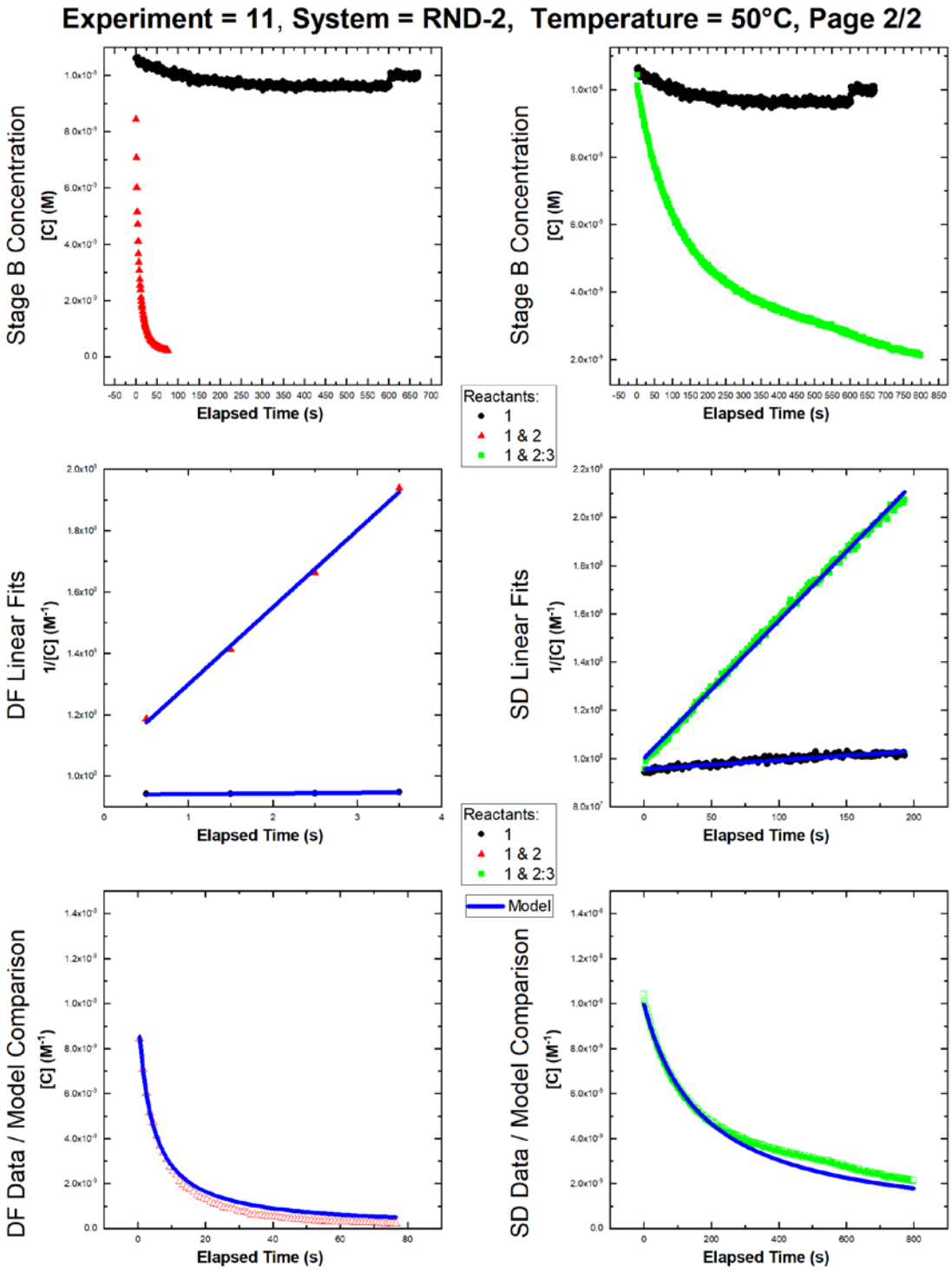

Experiment = 9, System = RND-2, Temperature = 30°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

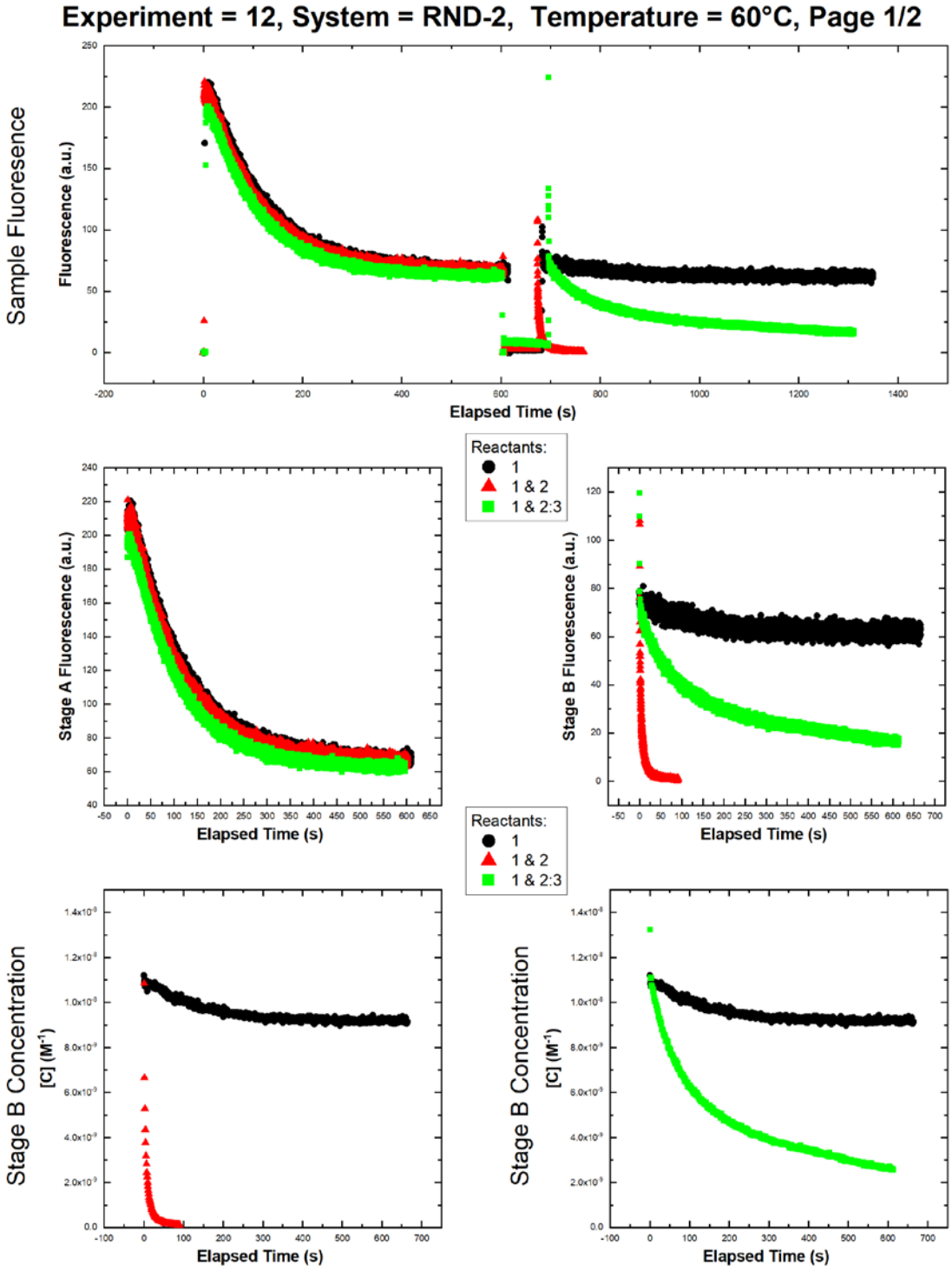

Experiment = 9, System = RND-2, Temperature = 30°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

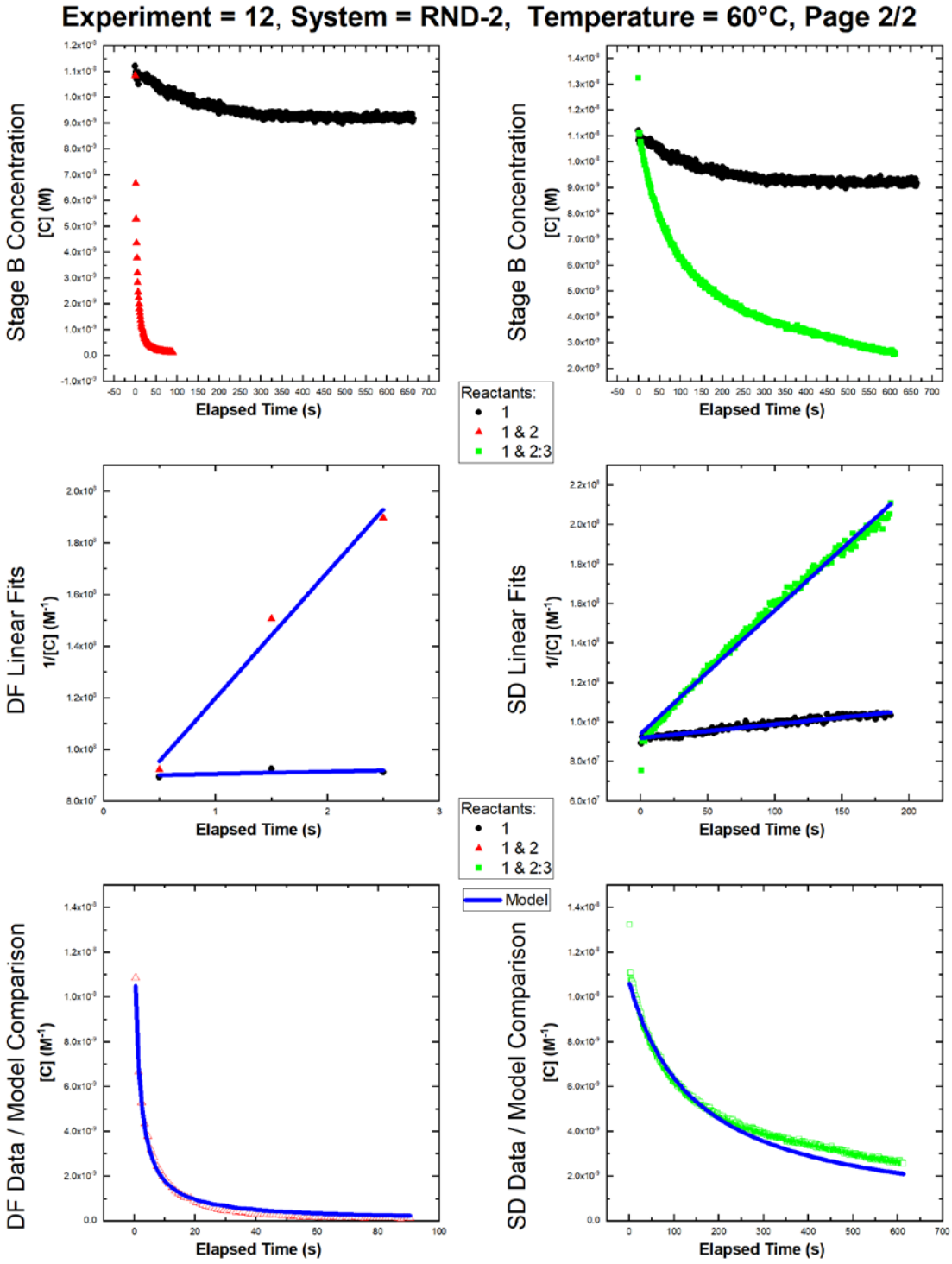

Experiment = 10, System = RND-2, Temperature = 40°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

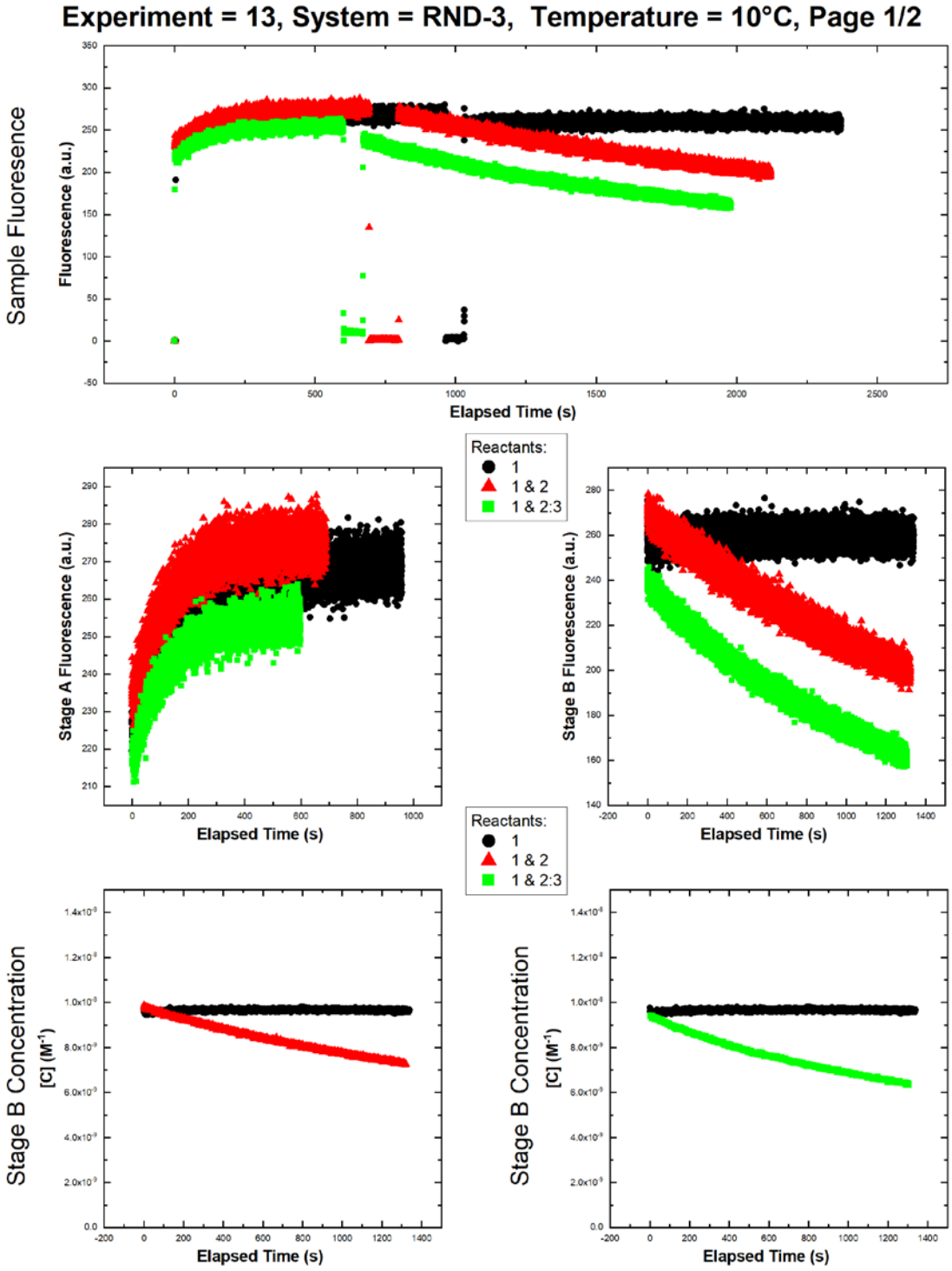

Experiment = 10, System = RND-2, Temperature = 40°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

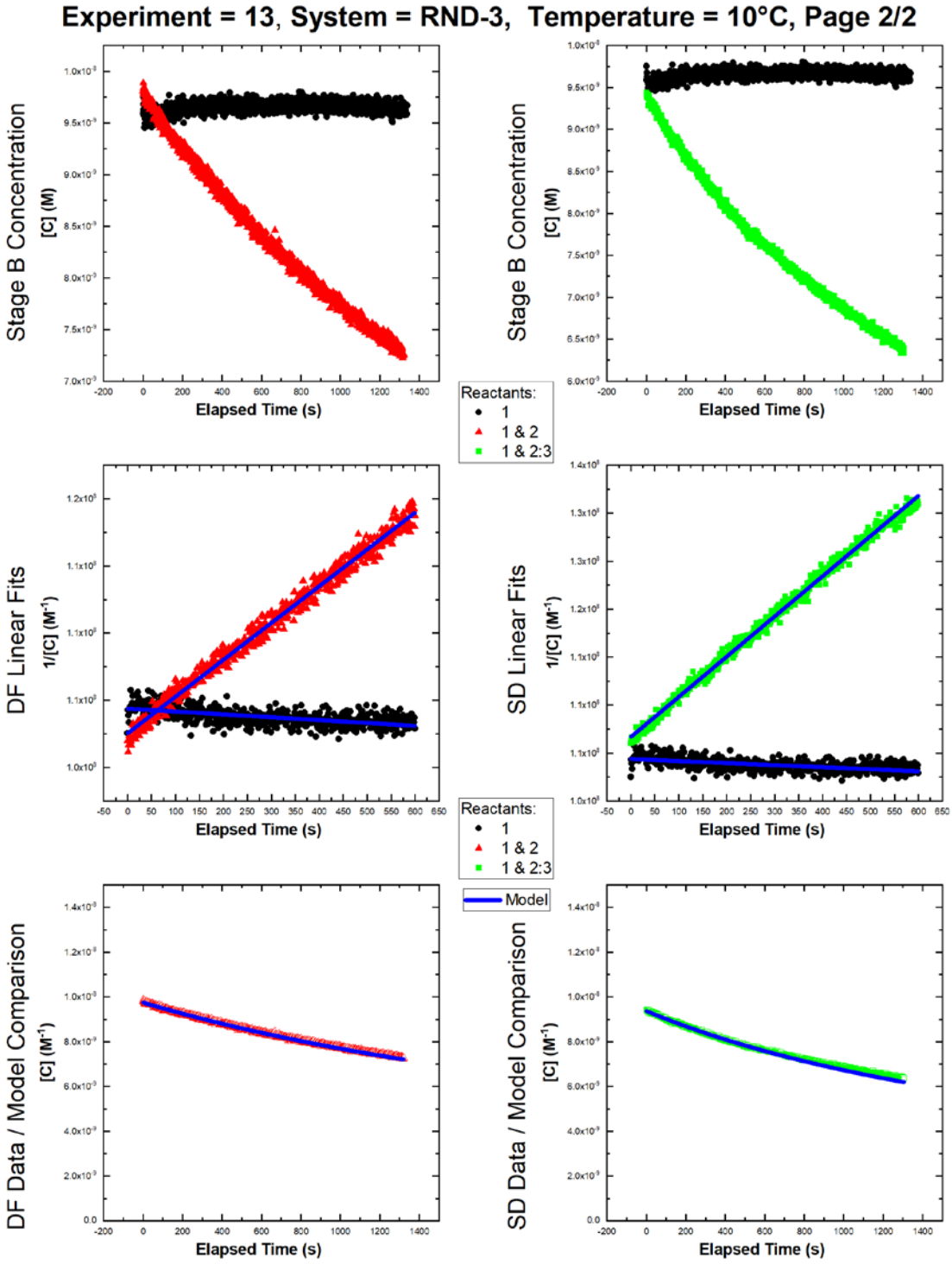

Experiment = 11, System = RND-2, Temperature = 50°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

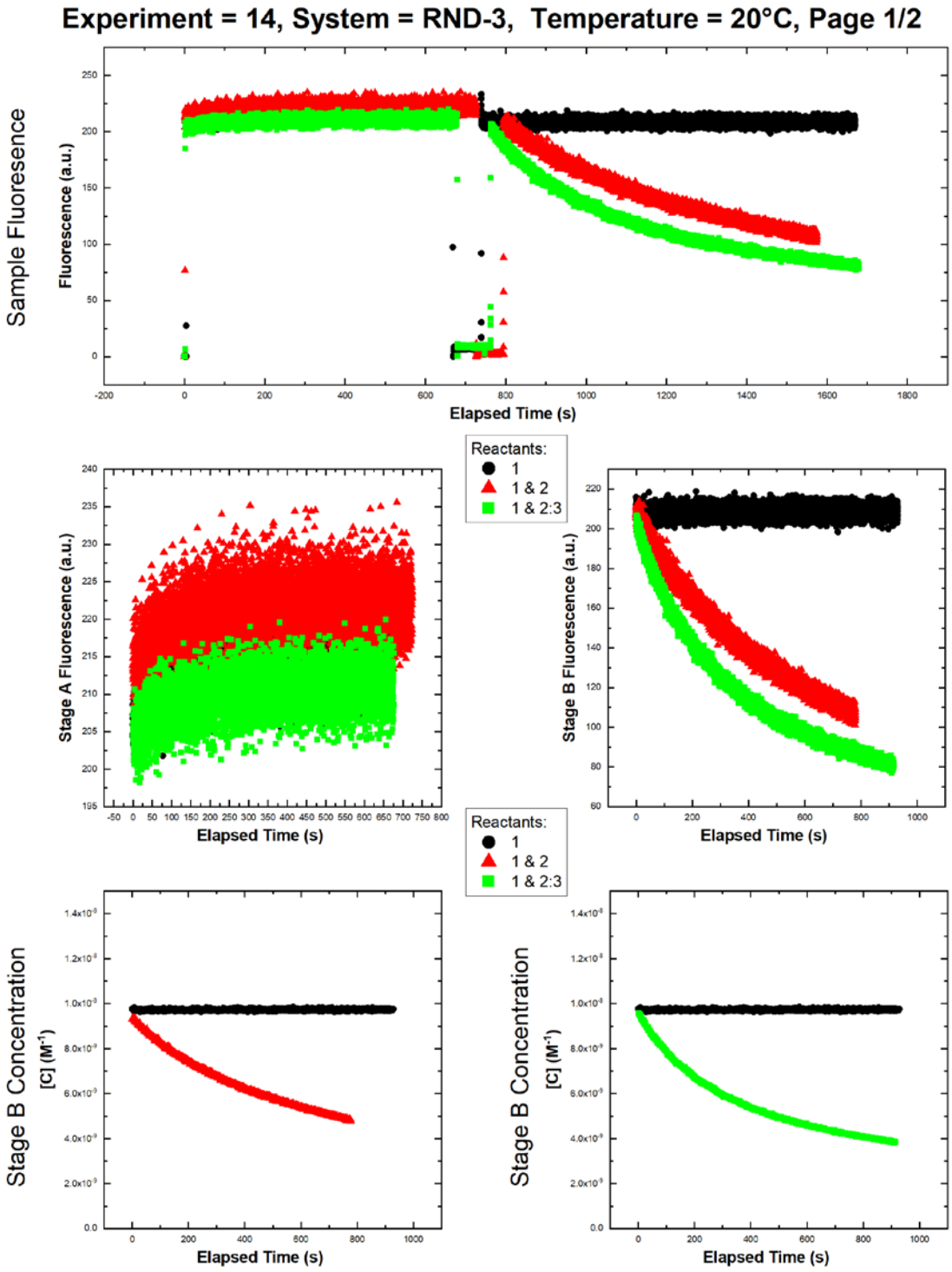

Experiment = 11, System = RND-2, Temperature = 50°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

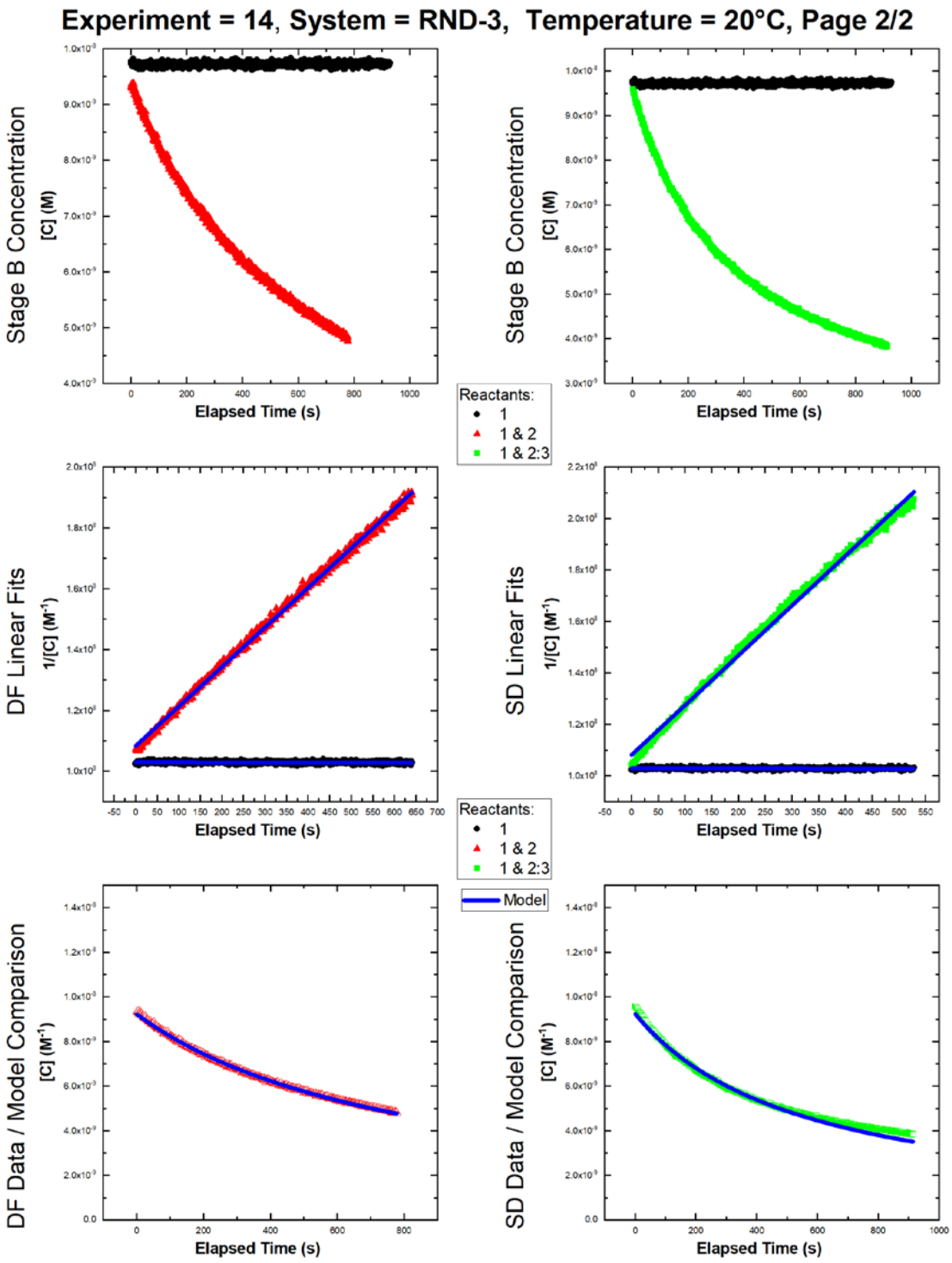

Experiment = 12, System = RND-2, Temperature = 60°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

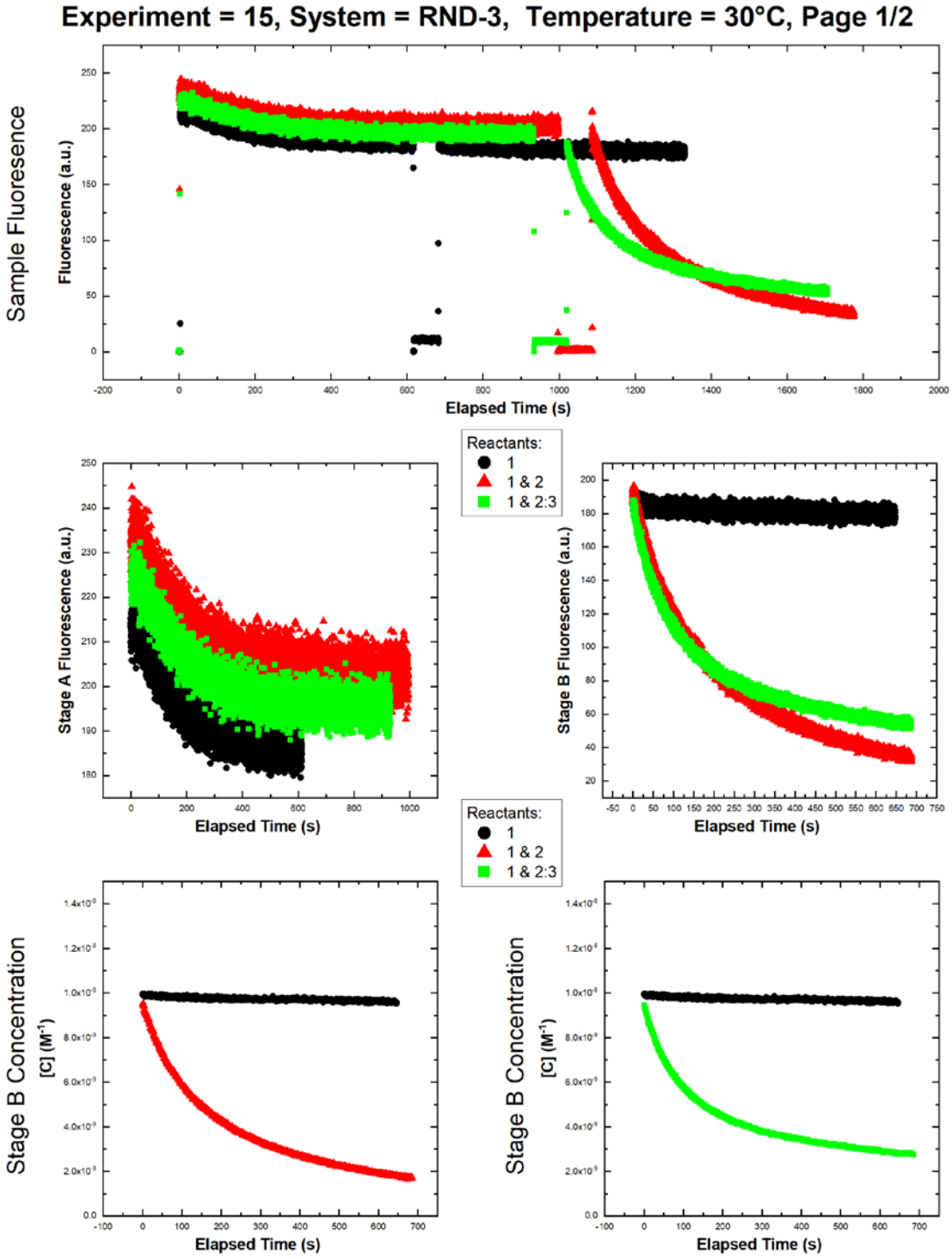

Experiment = 12, System = RND-2, Temperature = 60°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
[C] (M-1)
SD Data / Model Comparison

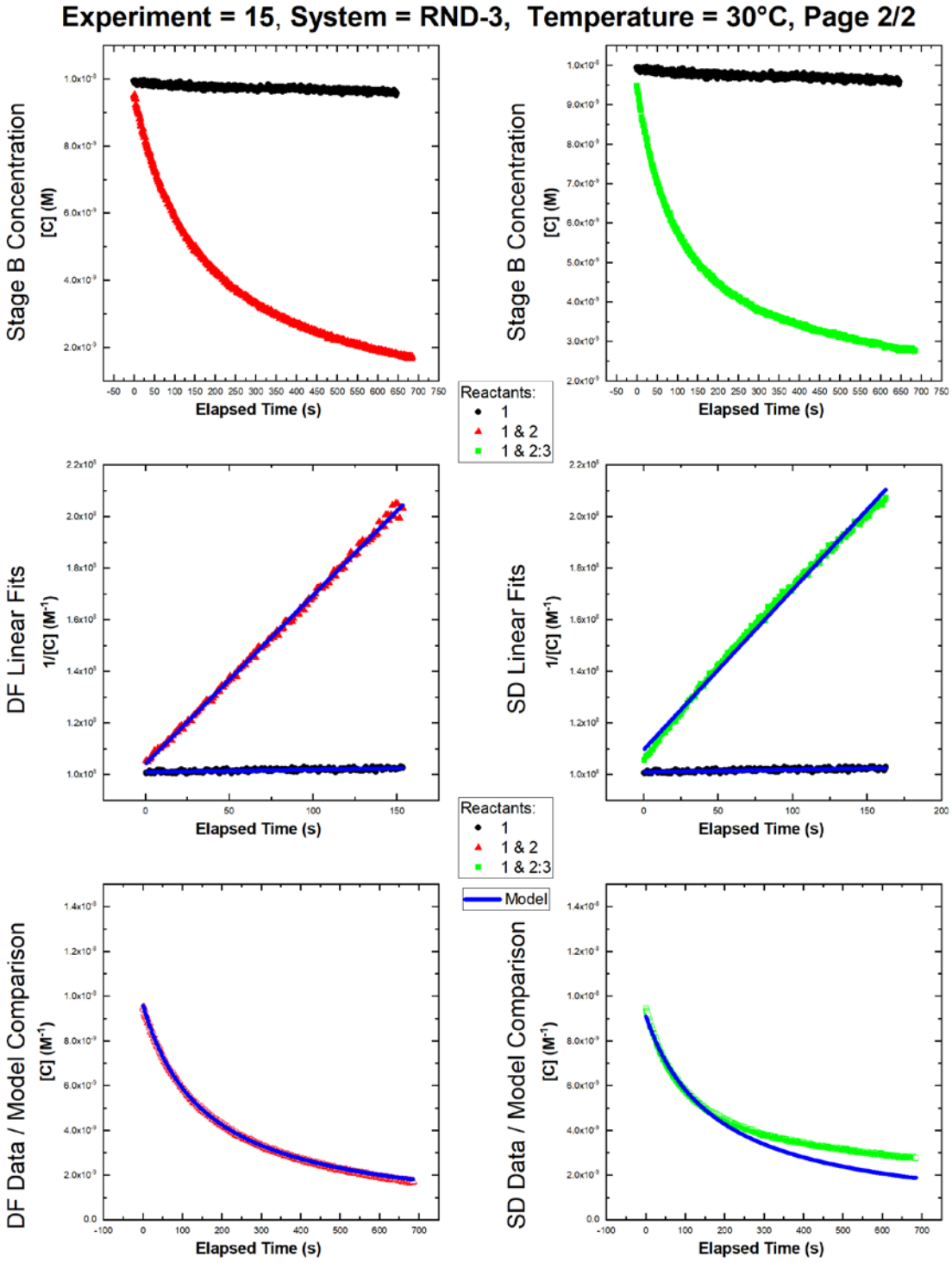

Experiment = 13, System = RND-3, Temperature = 10°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

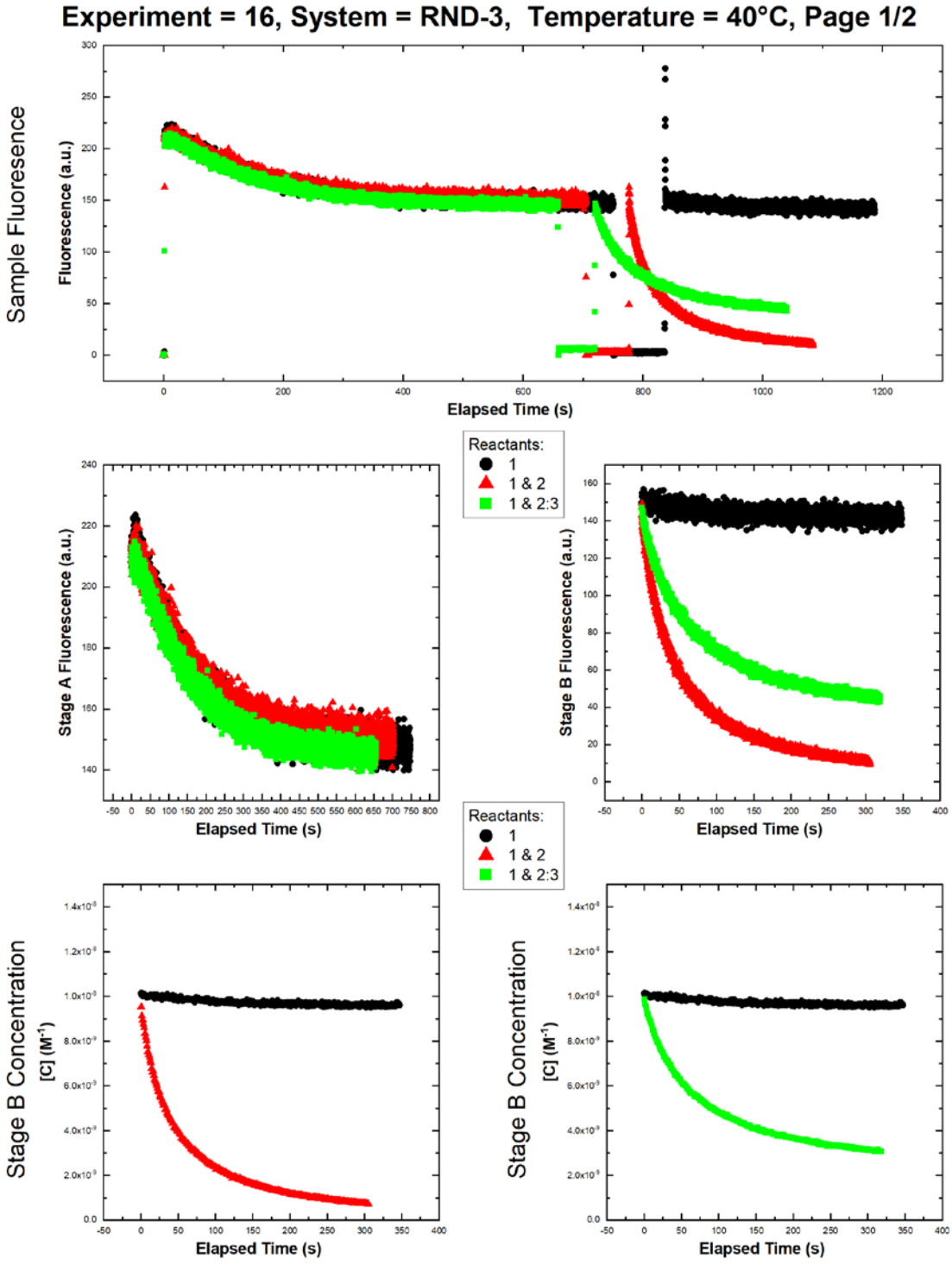

Experiment = 13, System = RND-3, Temperature = 10°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

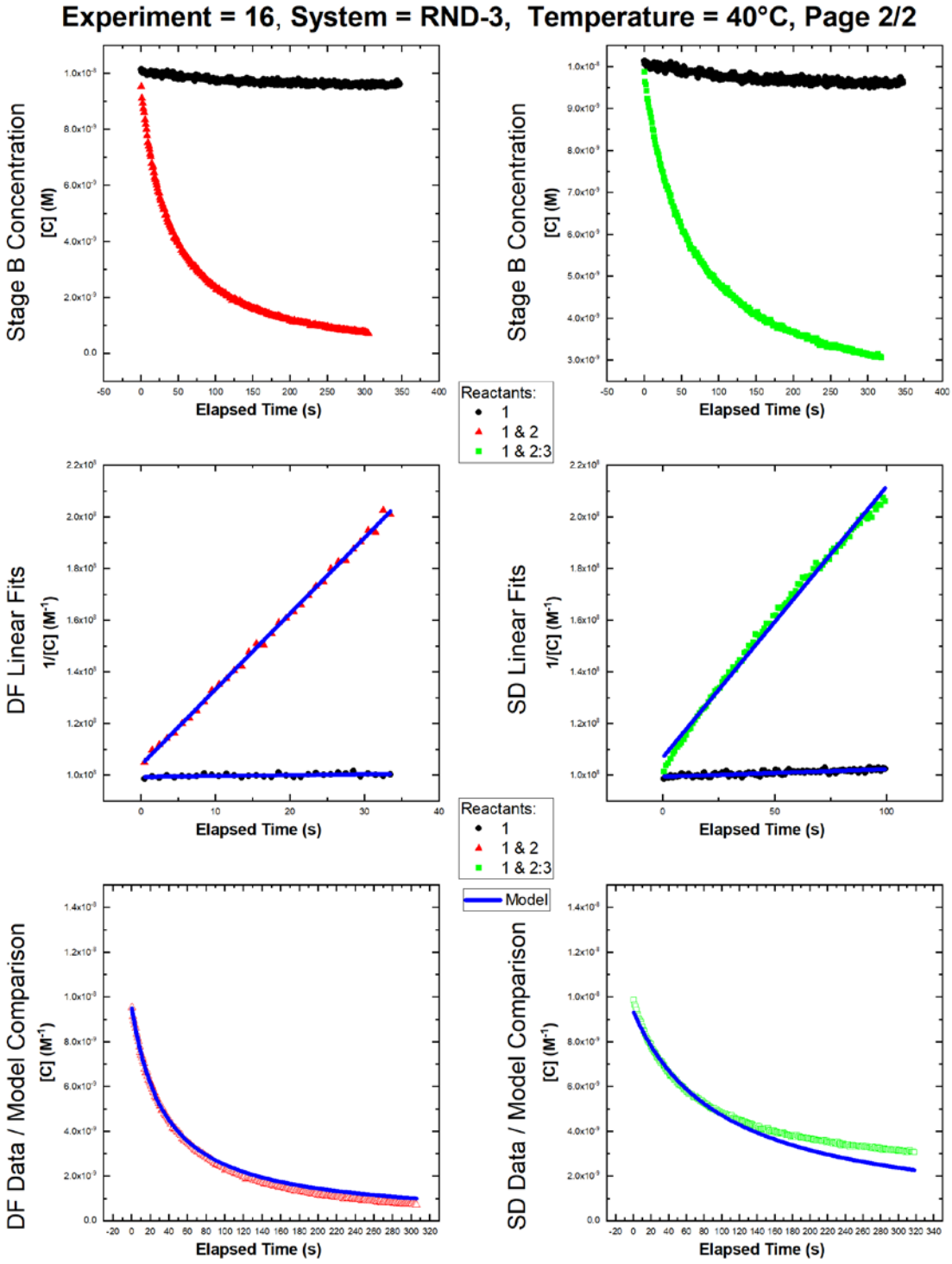

Experiment = 14, System = RND-3, Temperature = 20°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

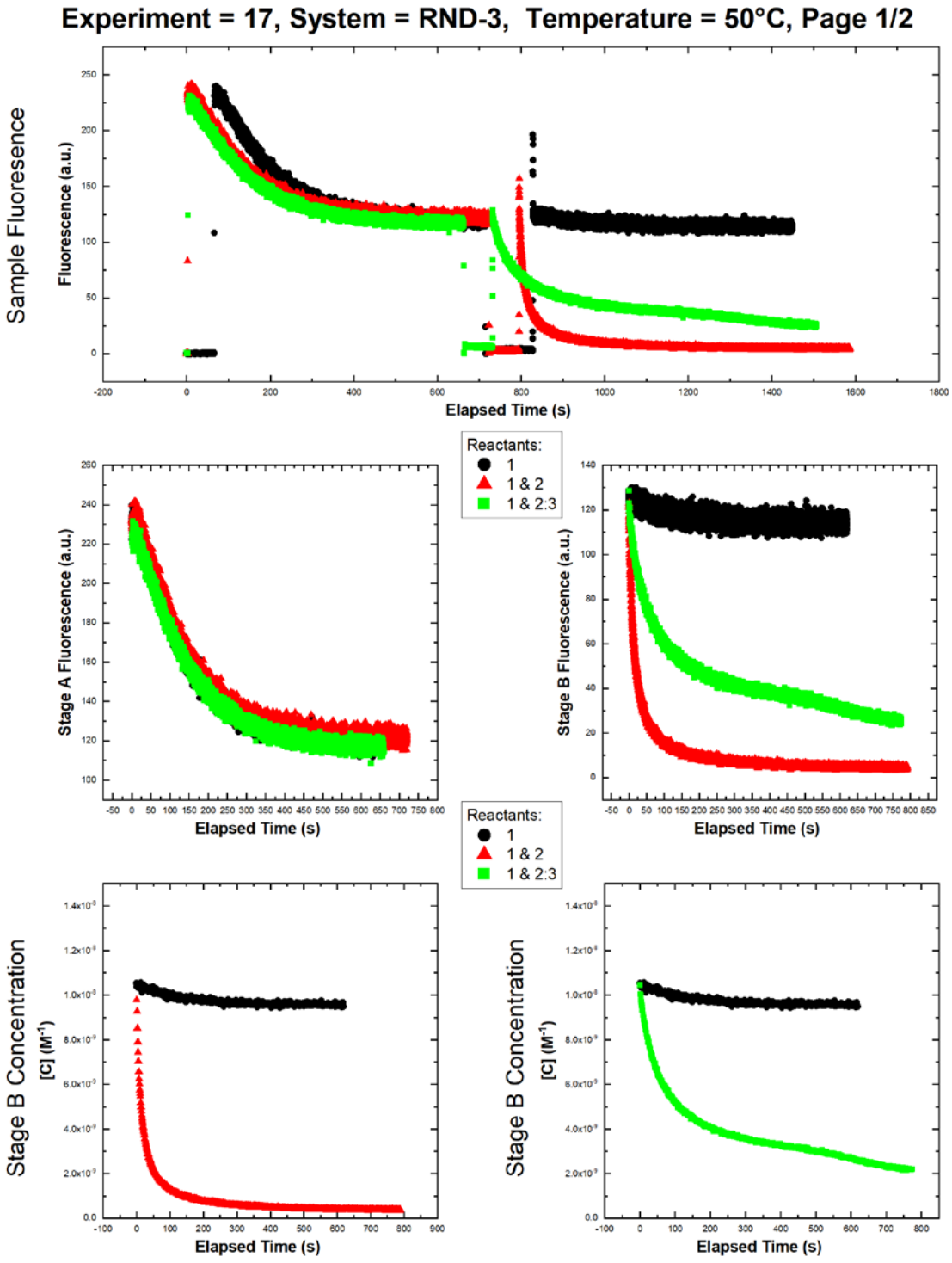

Experiment = 14, System = RND-3, Temperature = 20°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

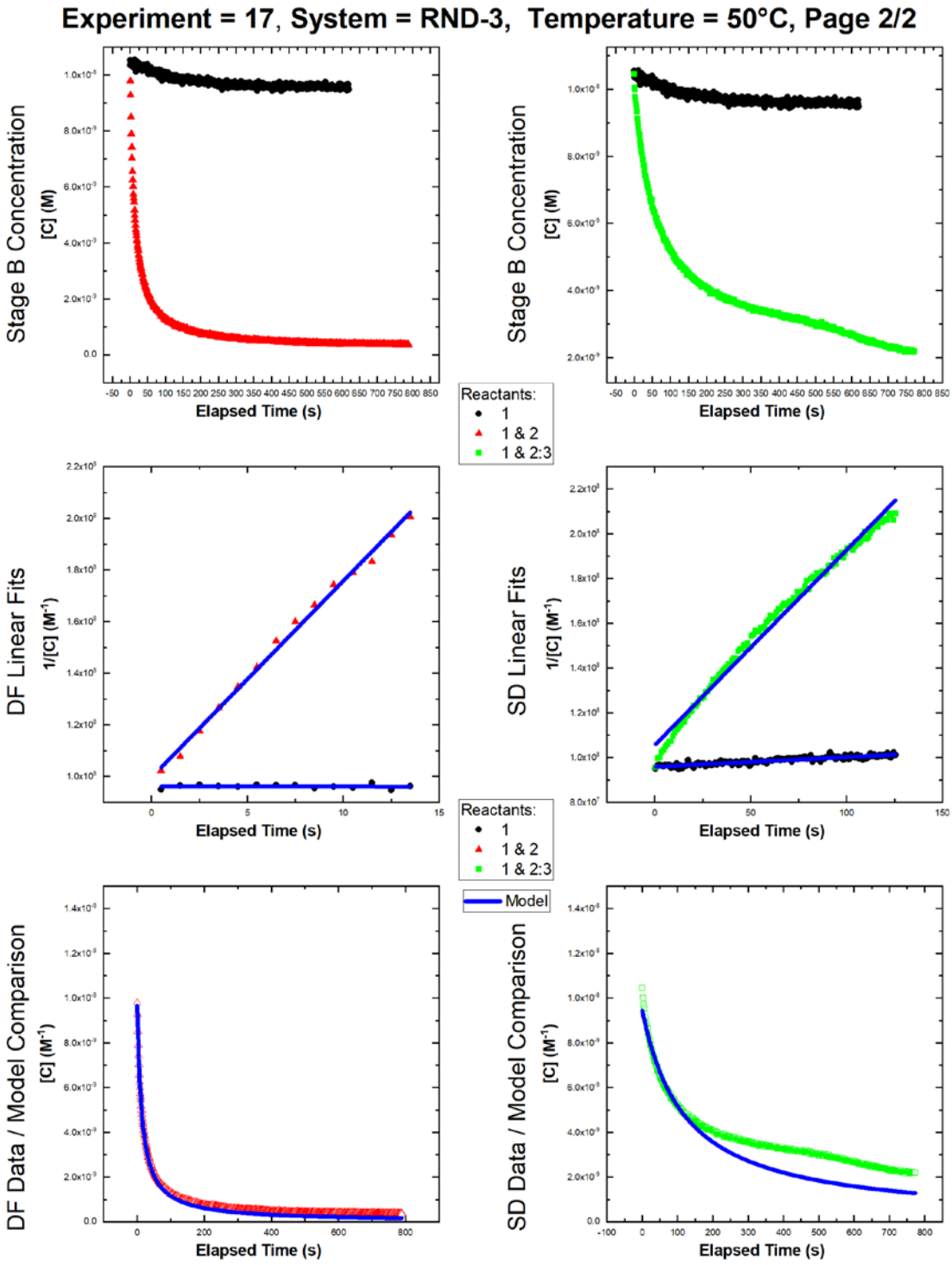

Experiment = 15, System = RND-3, Temperature = 30°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

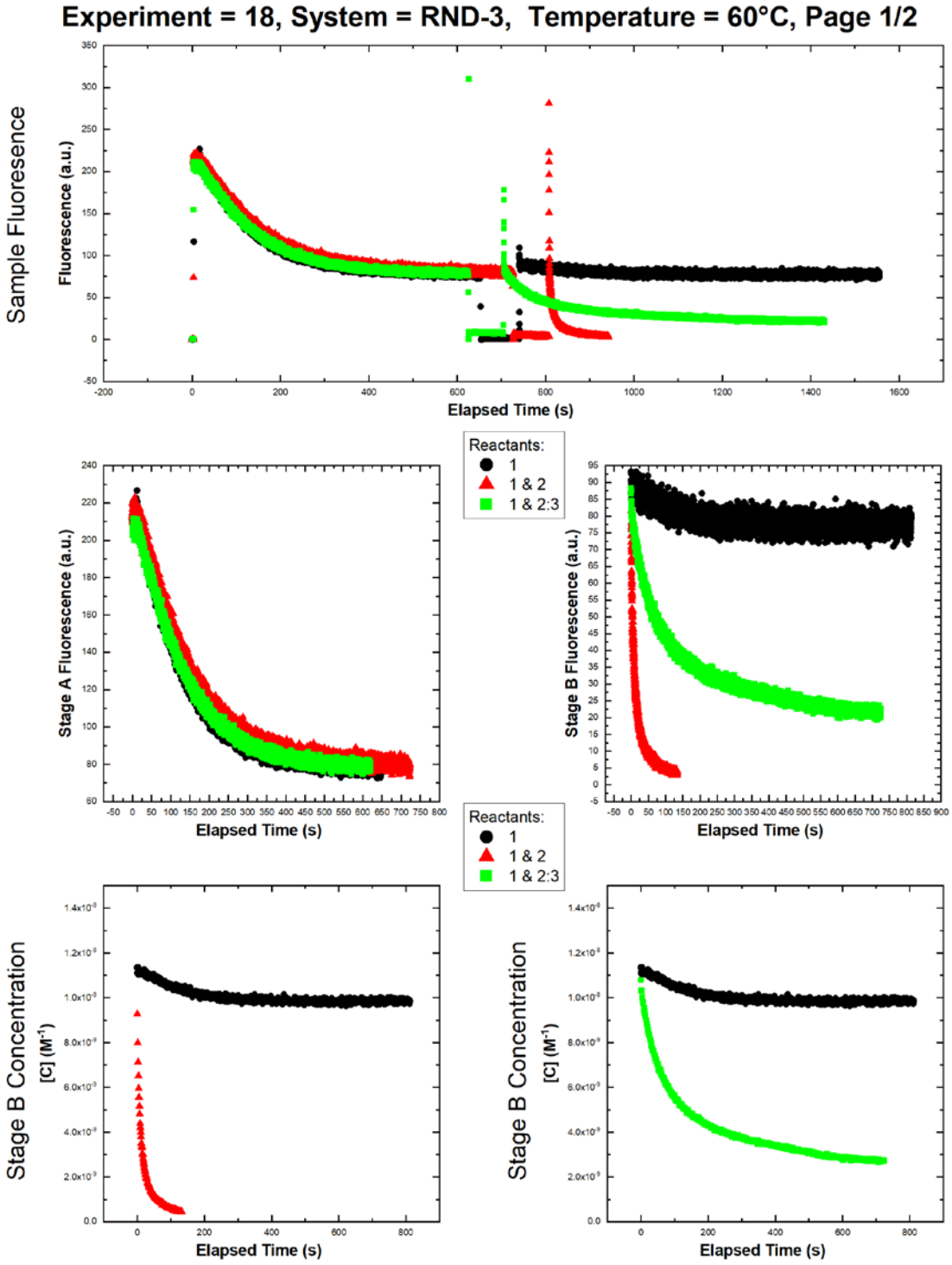

Experiment = 15, System = RND-3, Temperature = 30°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

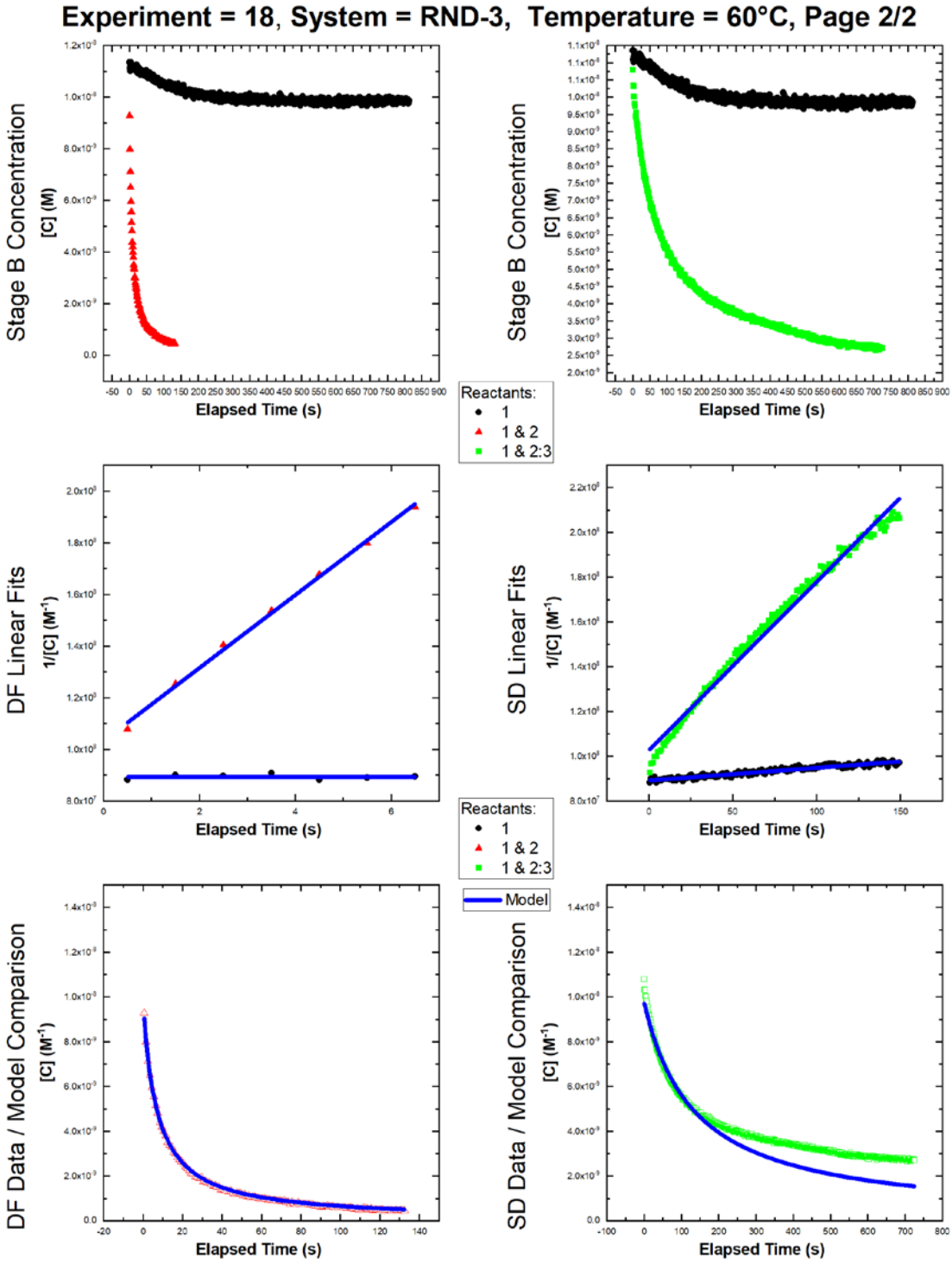

Experiment = 16, System = RND-3, Temperature = 40°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

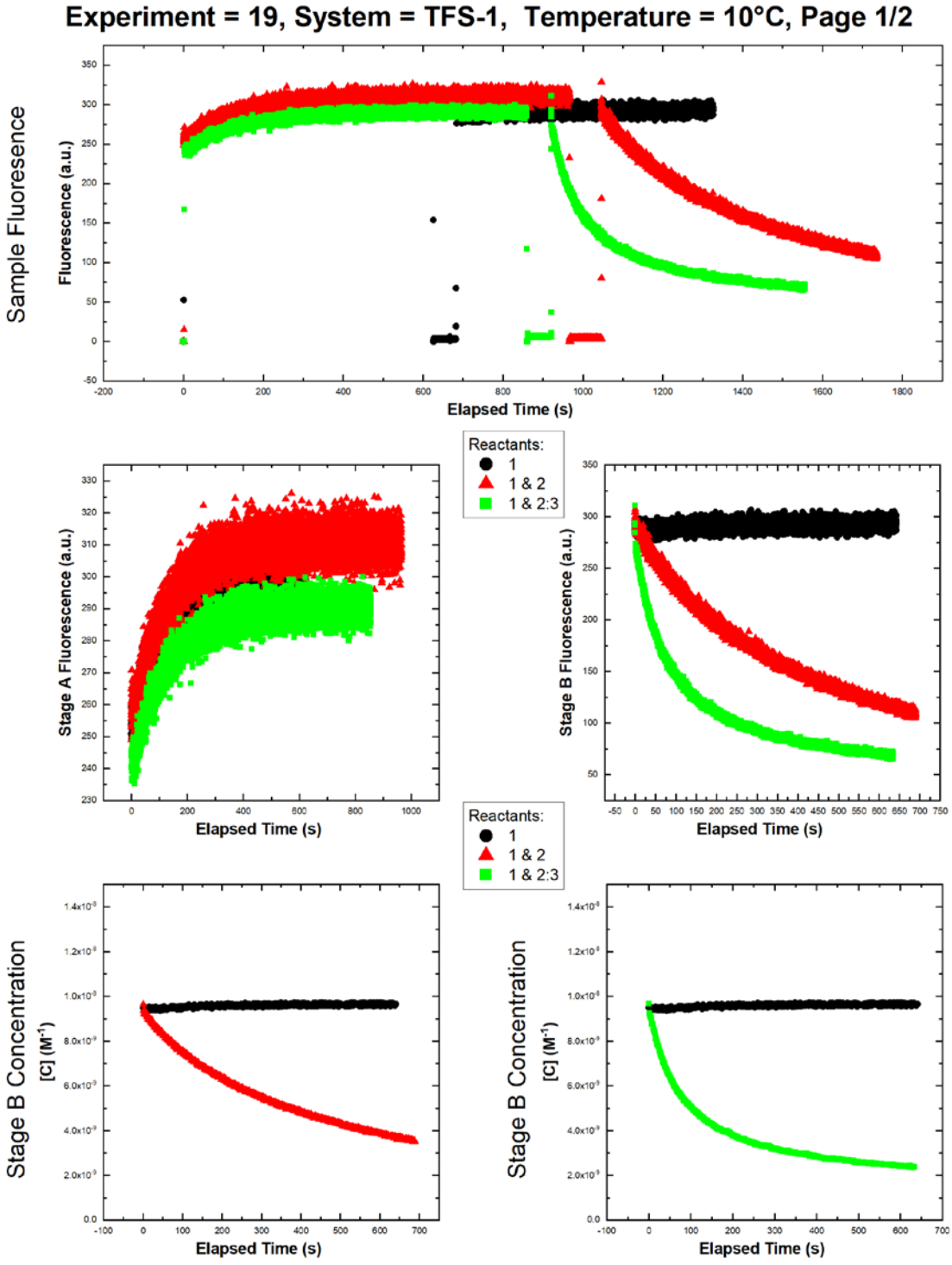

Experiment = 16, System = RND-3, Temperature = 40°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

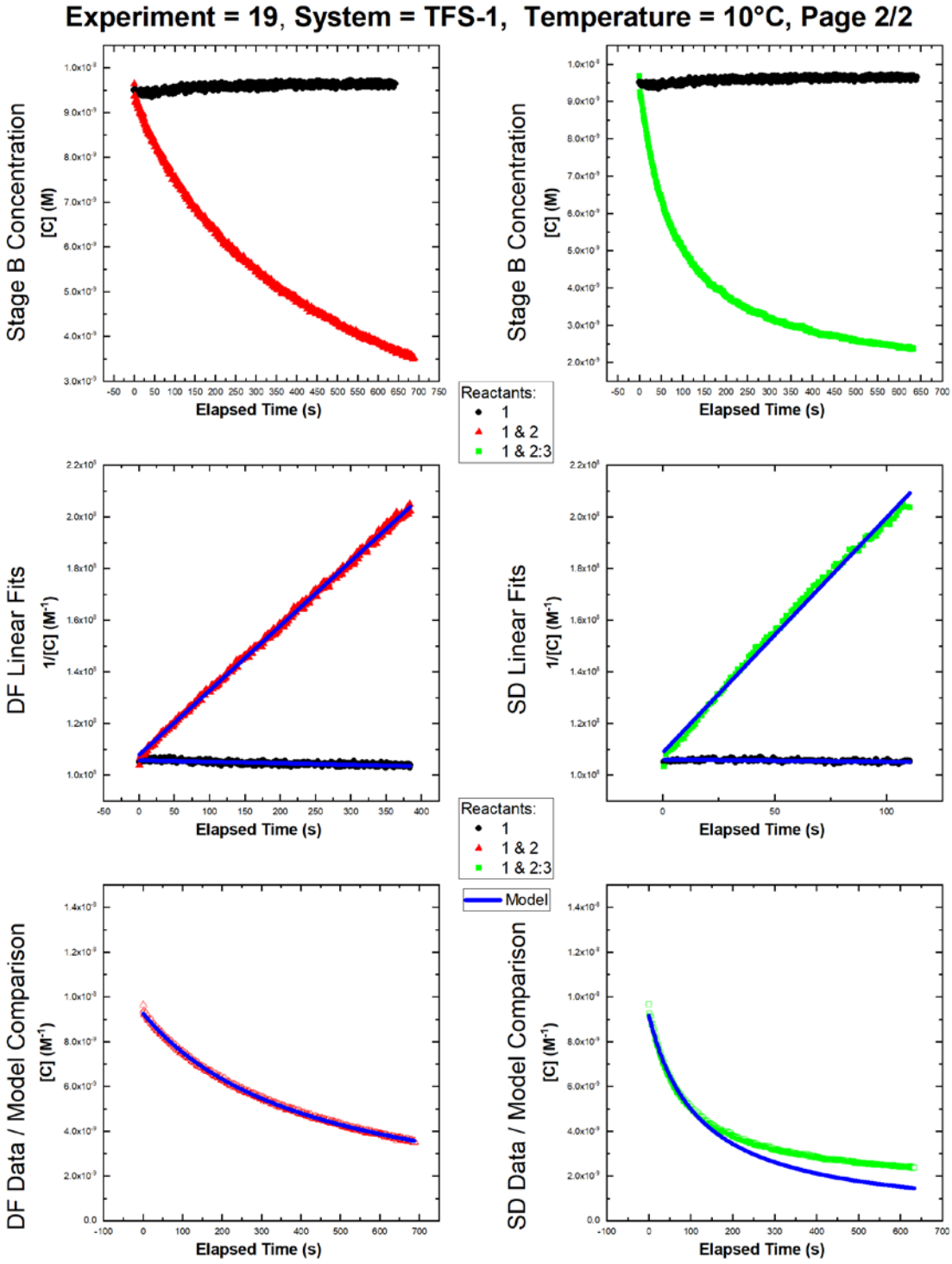

Experiment = 17, System = RND-3, Temperature = 50°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

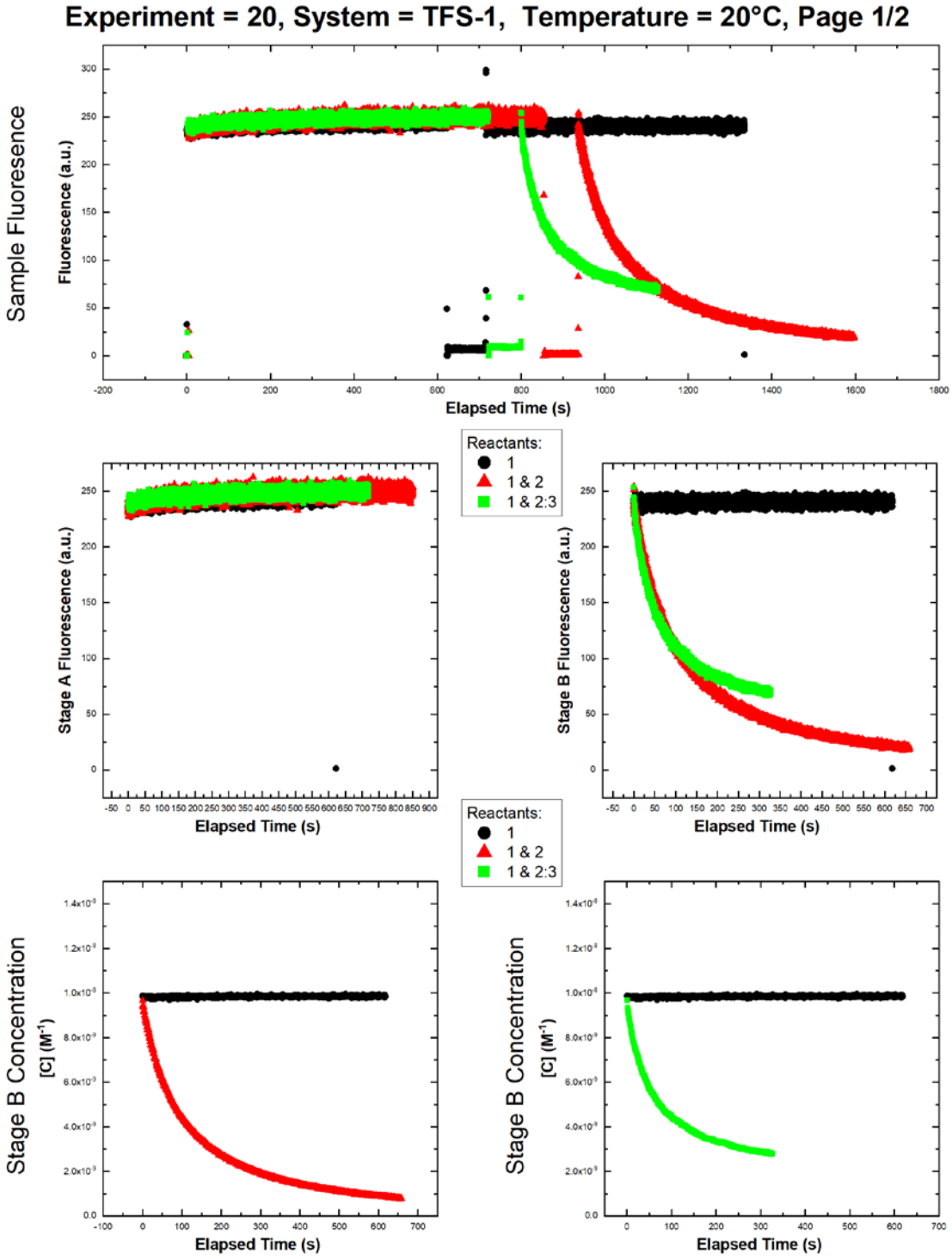

Experiment = 17, System = RND-3, Temperature = 50°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

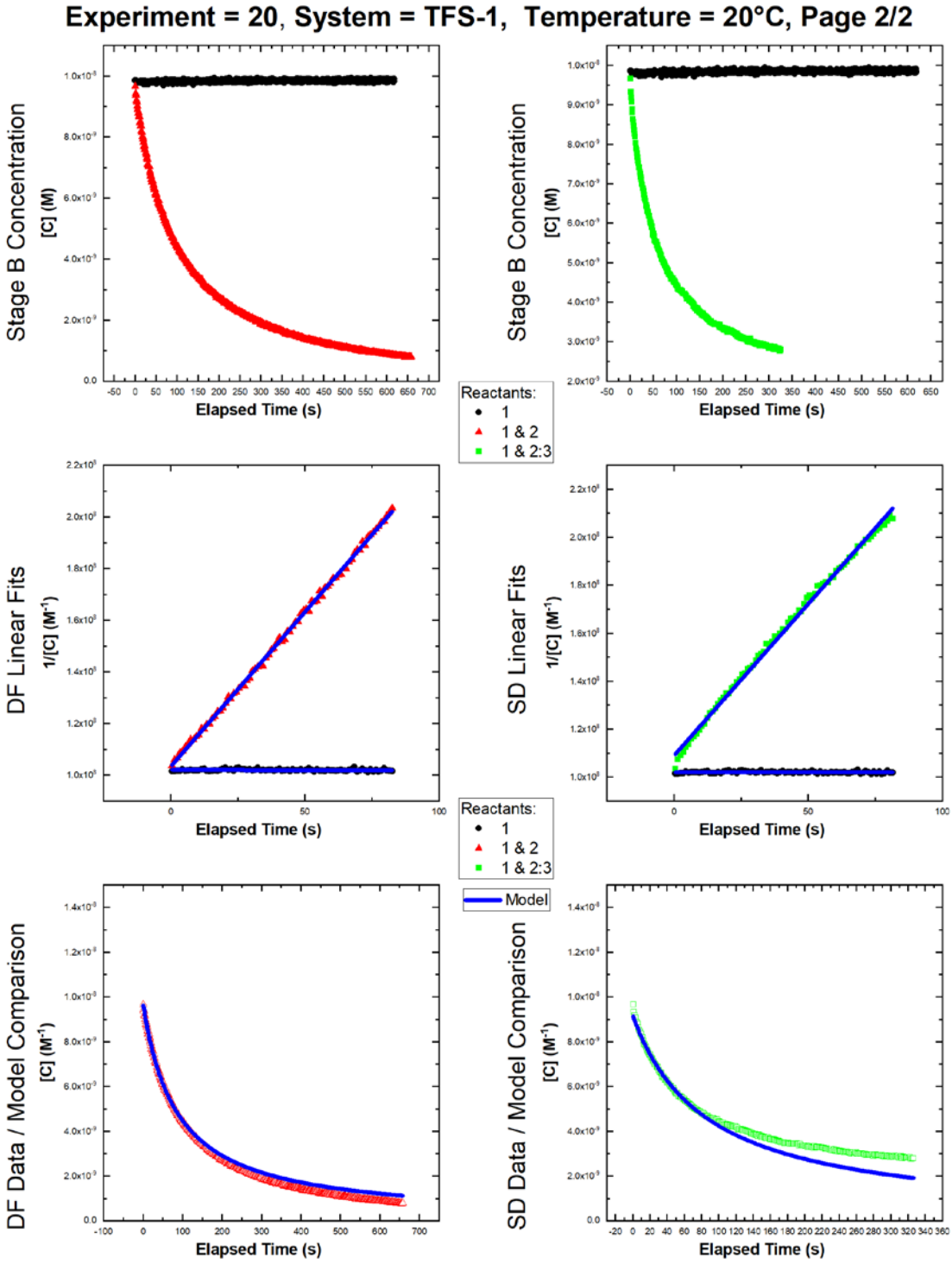

Experiment = 18, System = RND-3, Temperature = 60°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

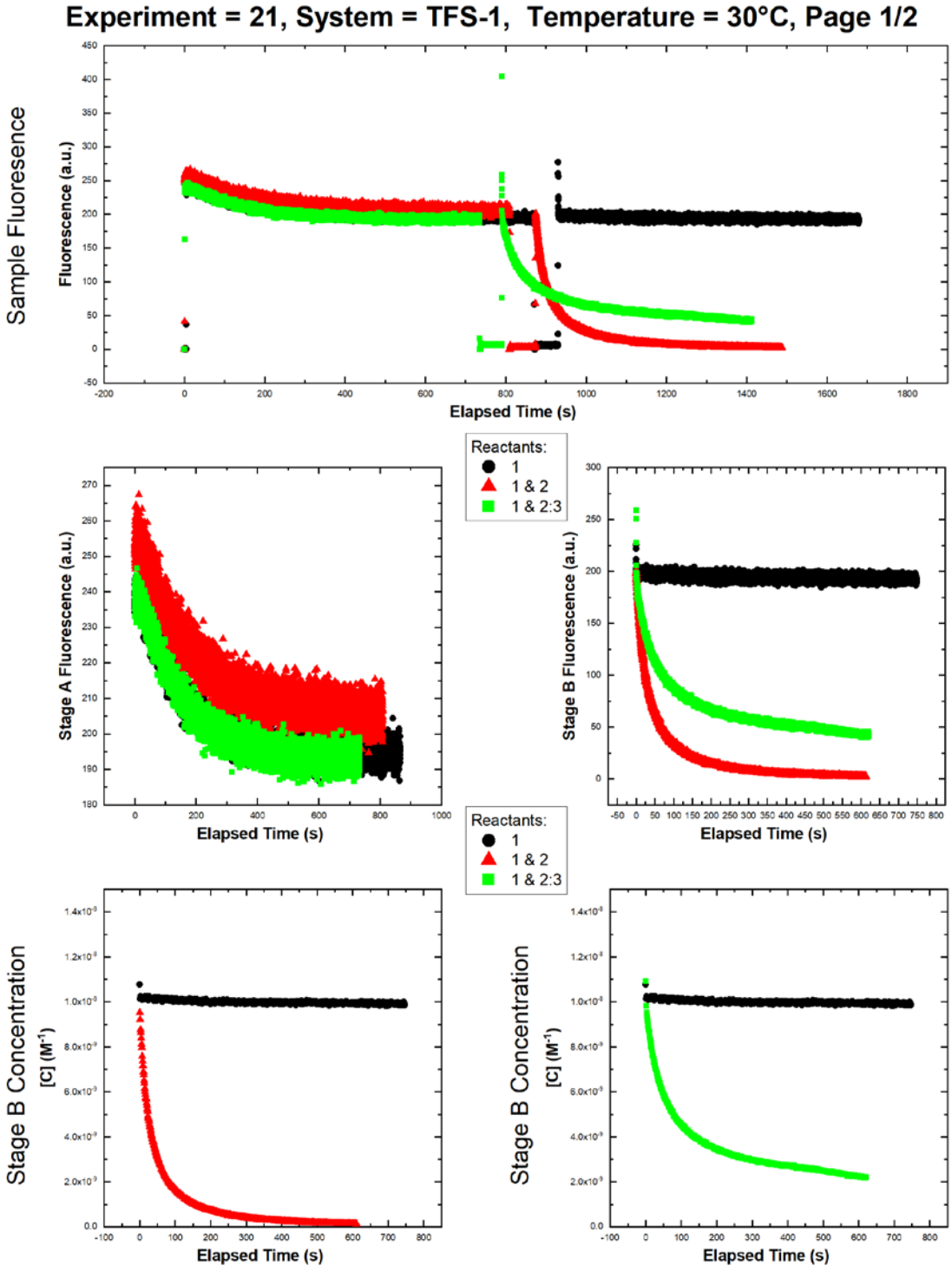

Experiment = 18, System = RND-3, Temperature = 60°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

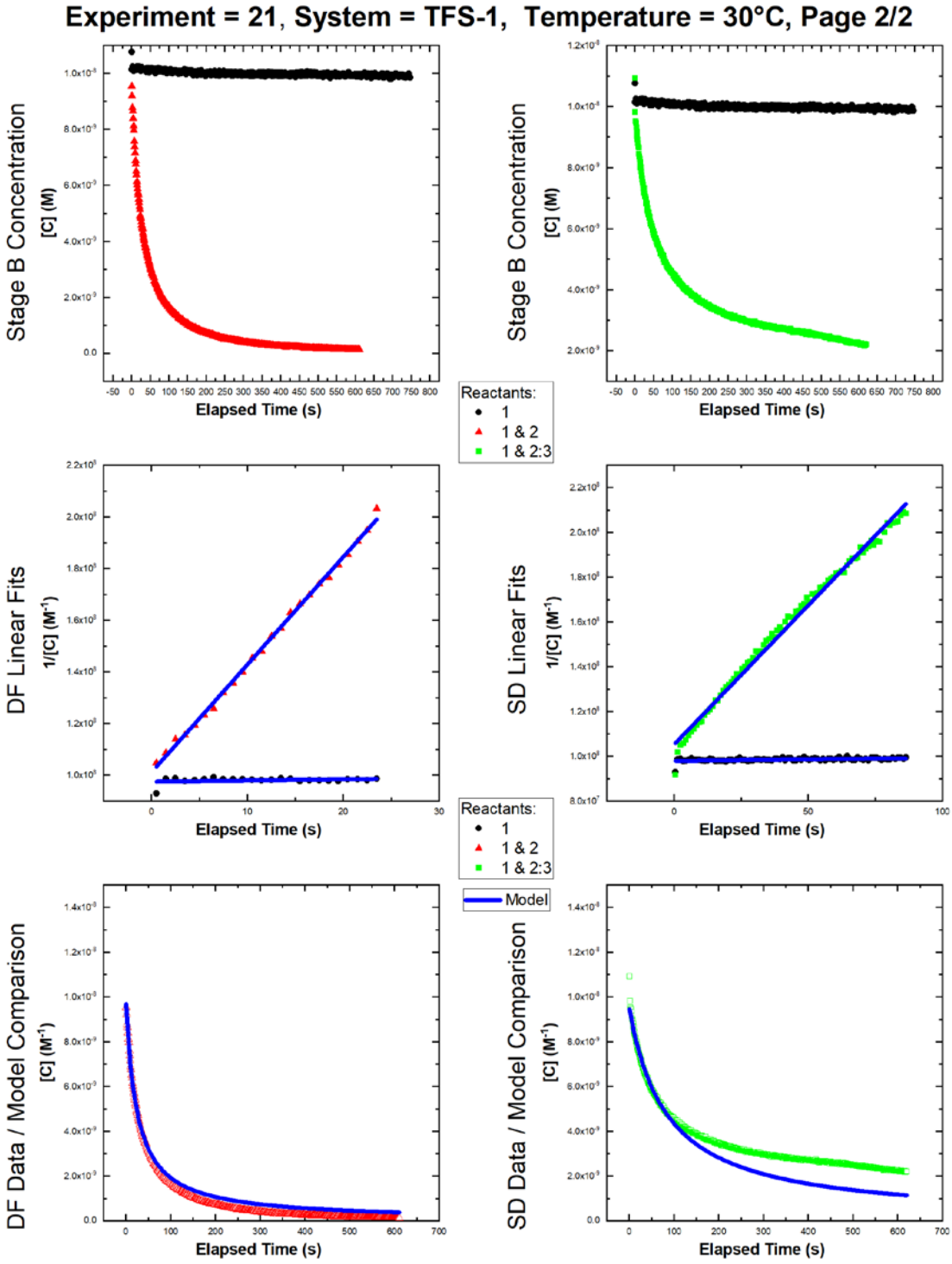

Experiment = 19, System = TFS-1, Temperature = 10°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

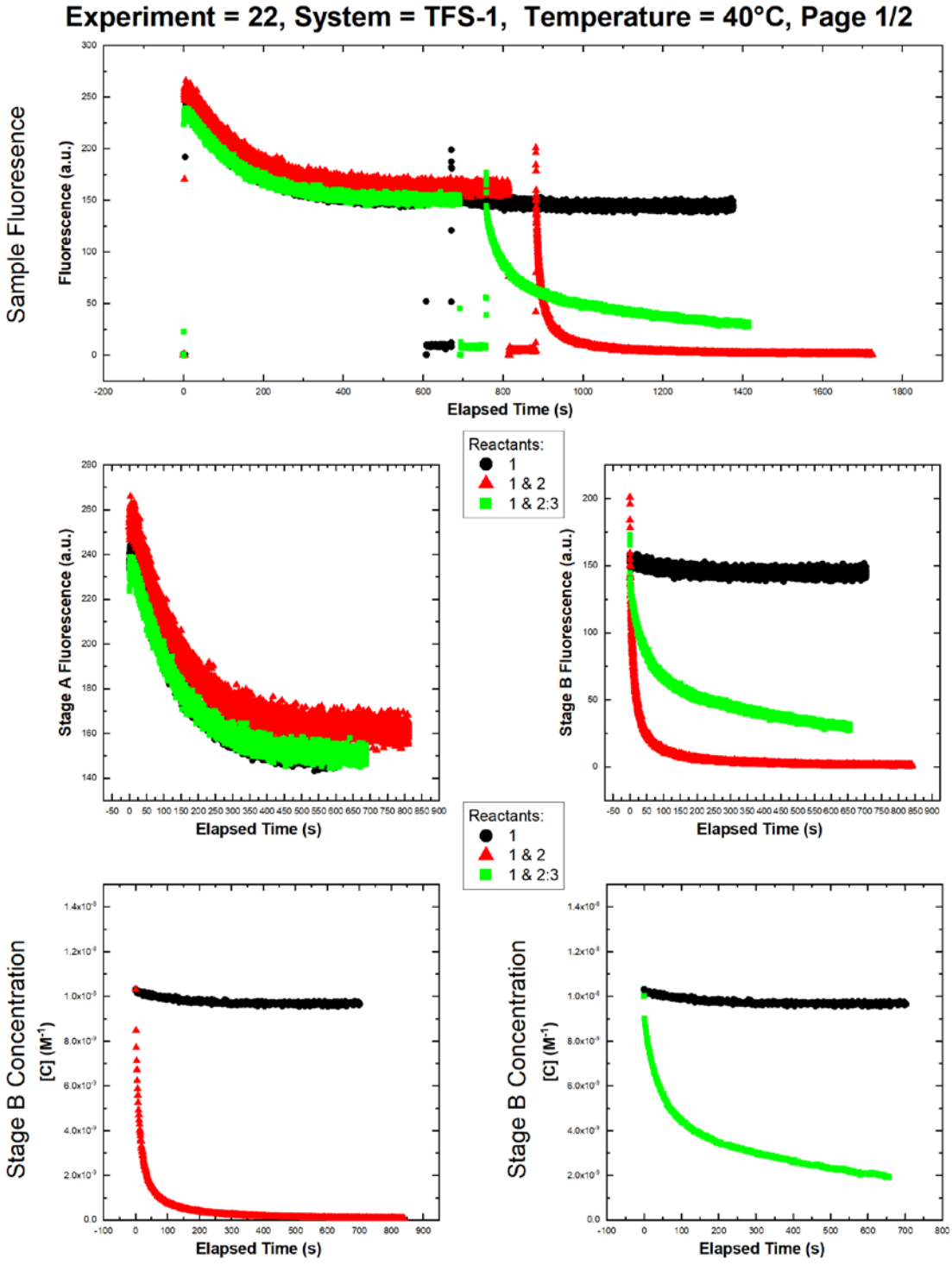

Experiment = 19, System = TFS-1, Temperature = 10°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

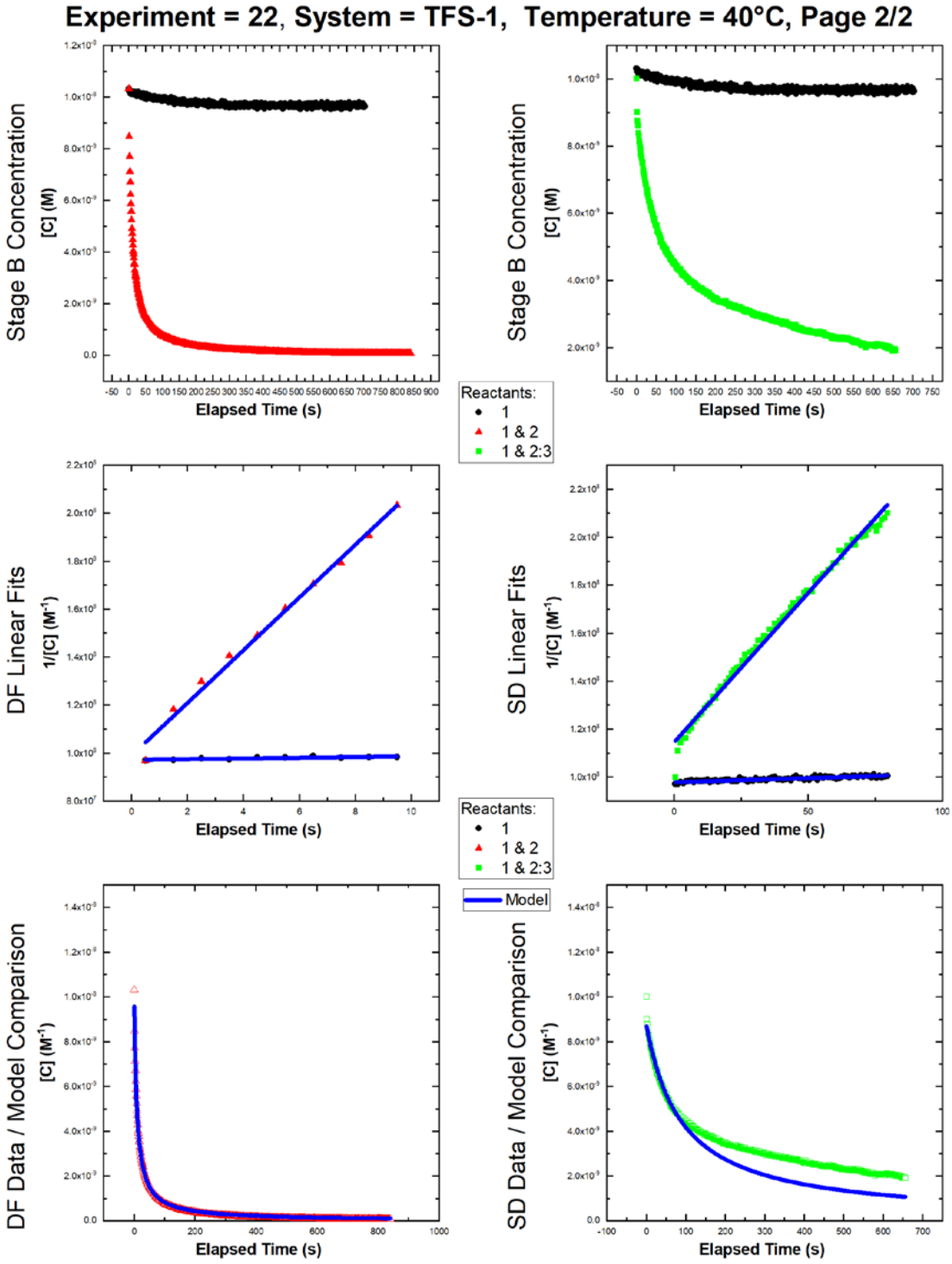

Experiment = 20, System = TFS-1, Temperature = 20°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Stage B Fluorescence (a.u.)
Stage B Concentration
[C] (M-1)

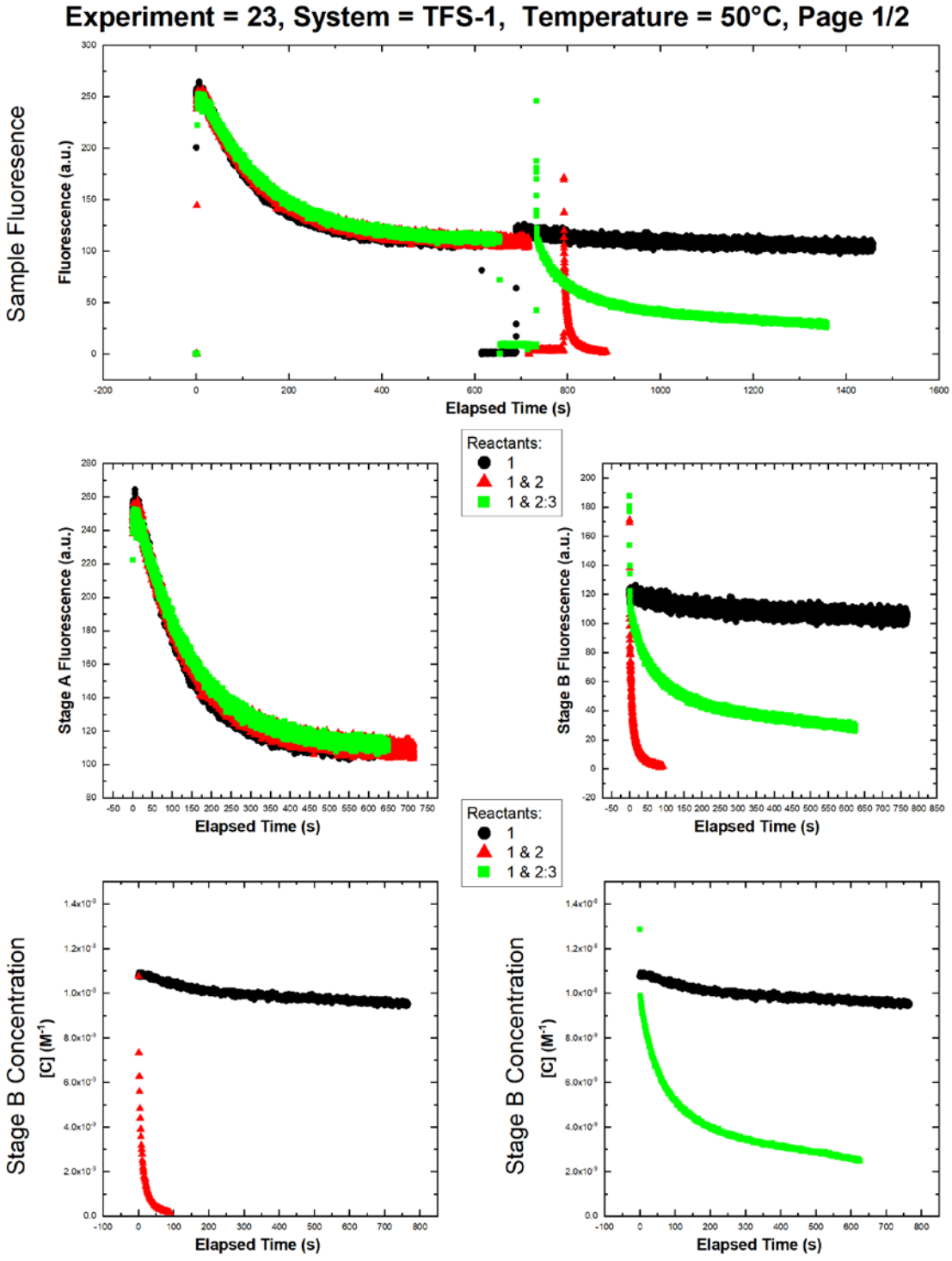

Experiment = 20, System = TFS-1, Temperature = 20°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

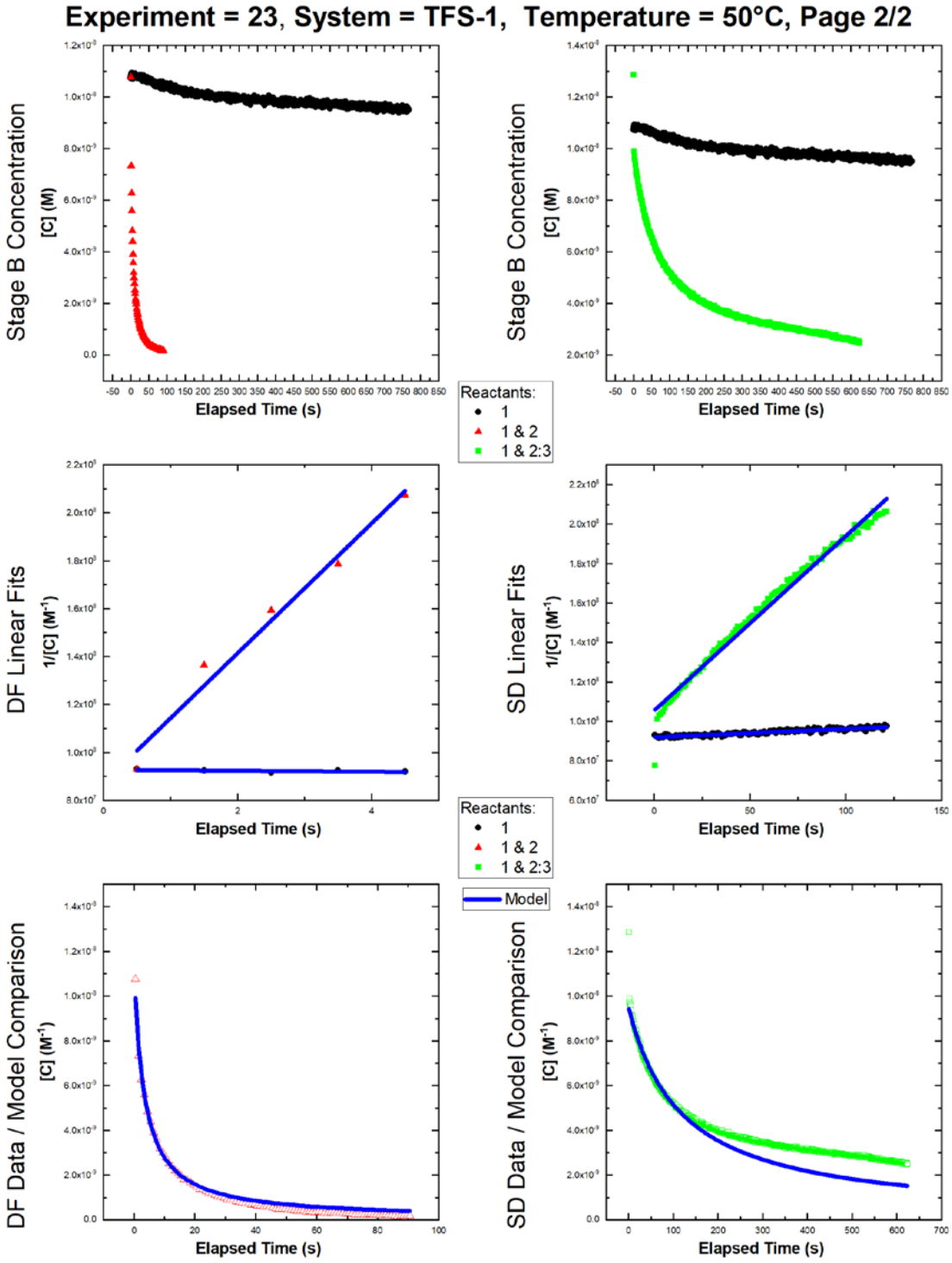

Experiment = 21, System = TFS-1, Temperature = 30°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

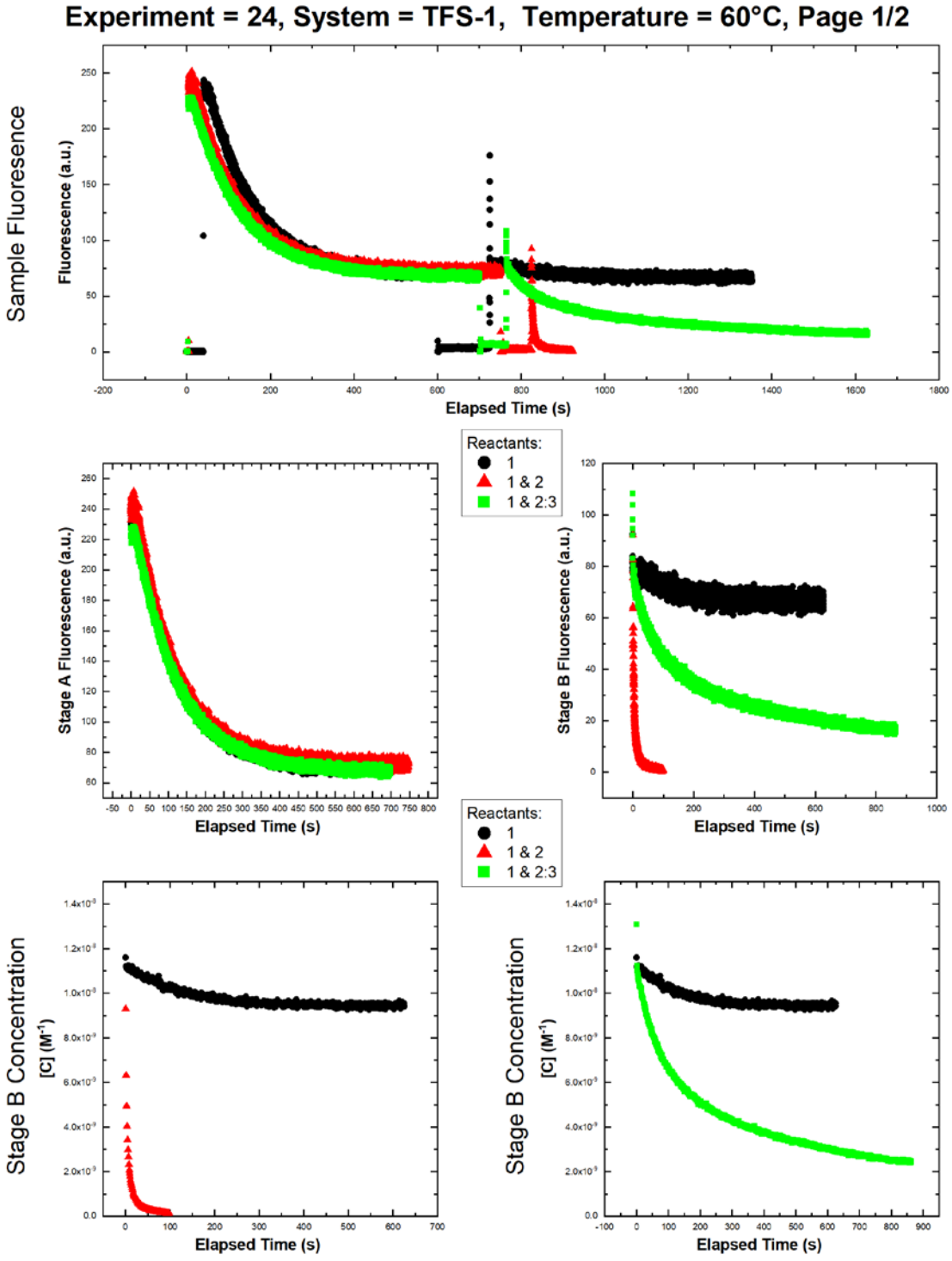

Experiment = 21, System = TFS-1, Temperature = 30°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

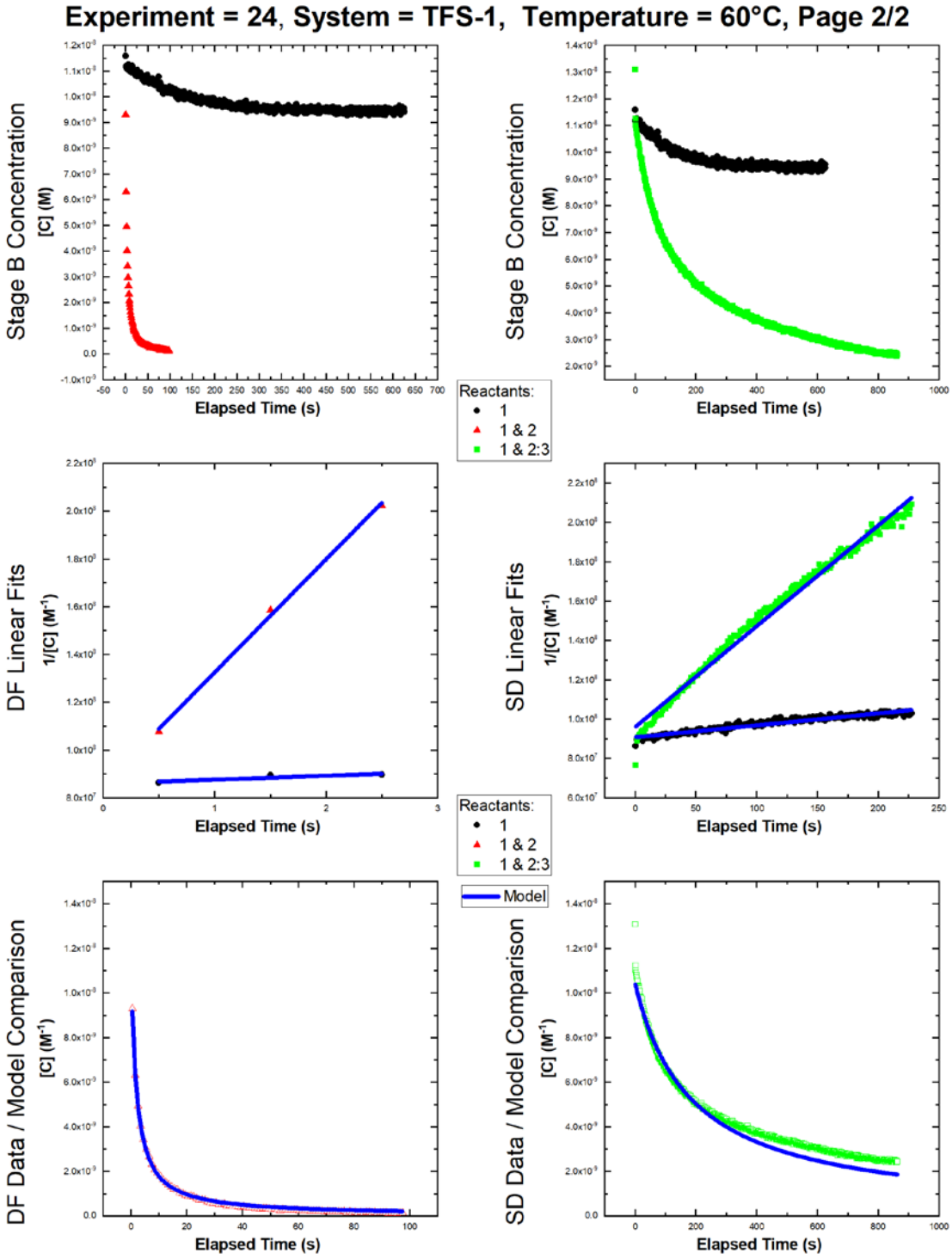

Experiment = 22, System = TFS-1, Temperature = 40°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

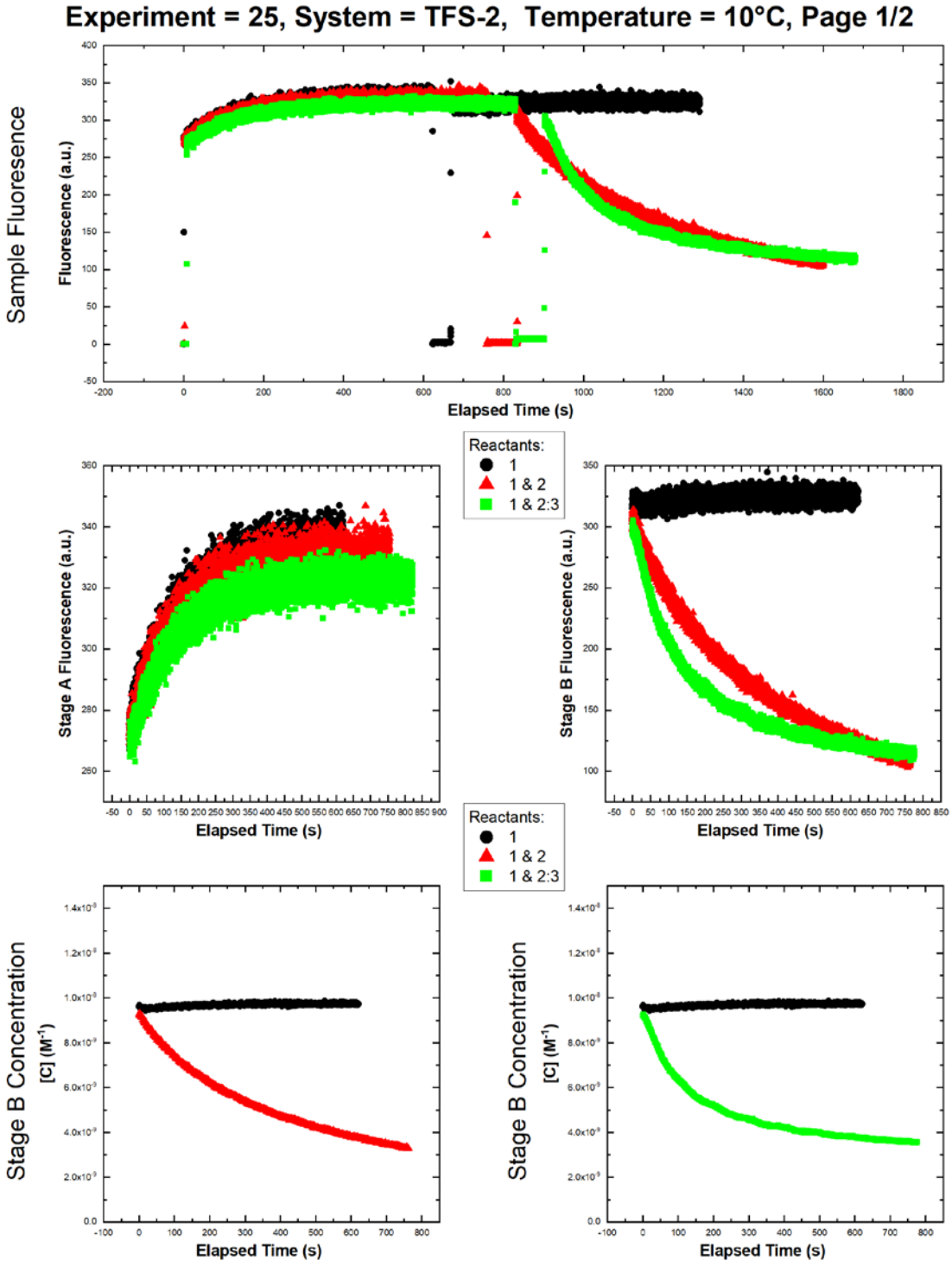

Experiment = 22, System = TFS-1, Temperature = 40°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

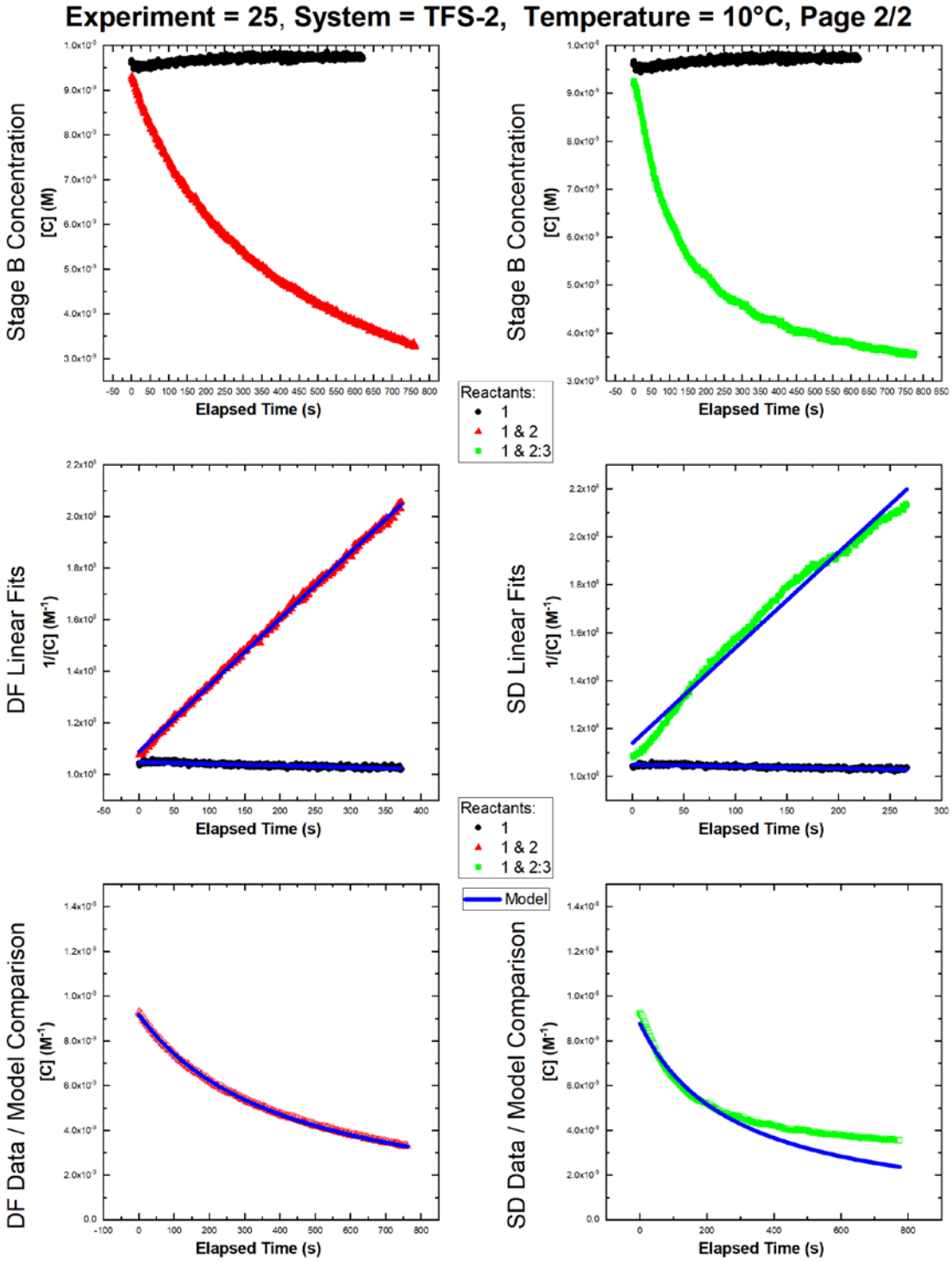

Experiment = 23, System = TFS-1, Temperature = 50°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

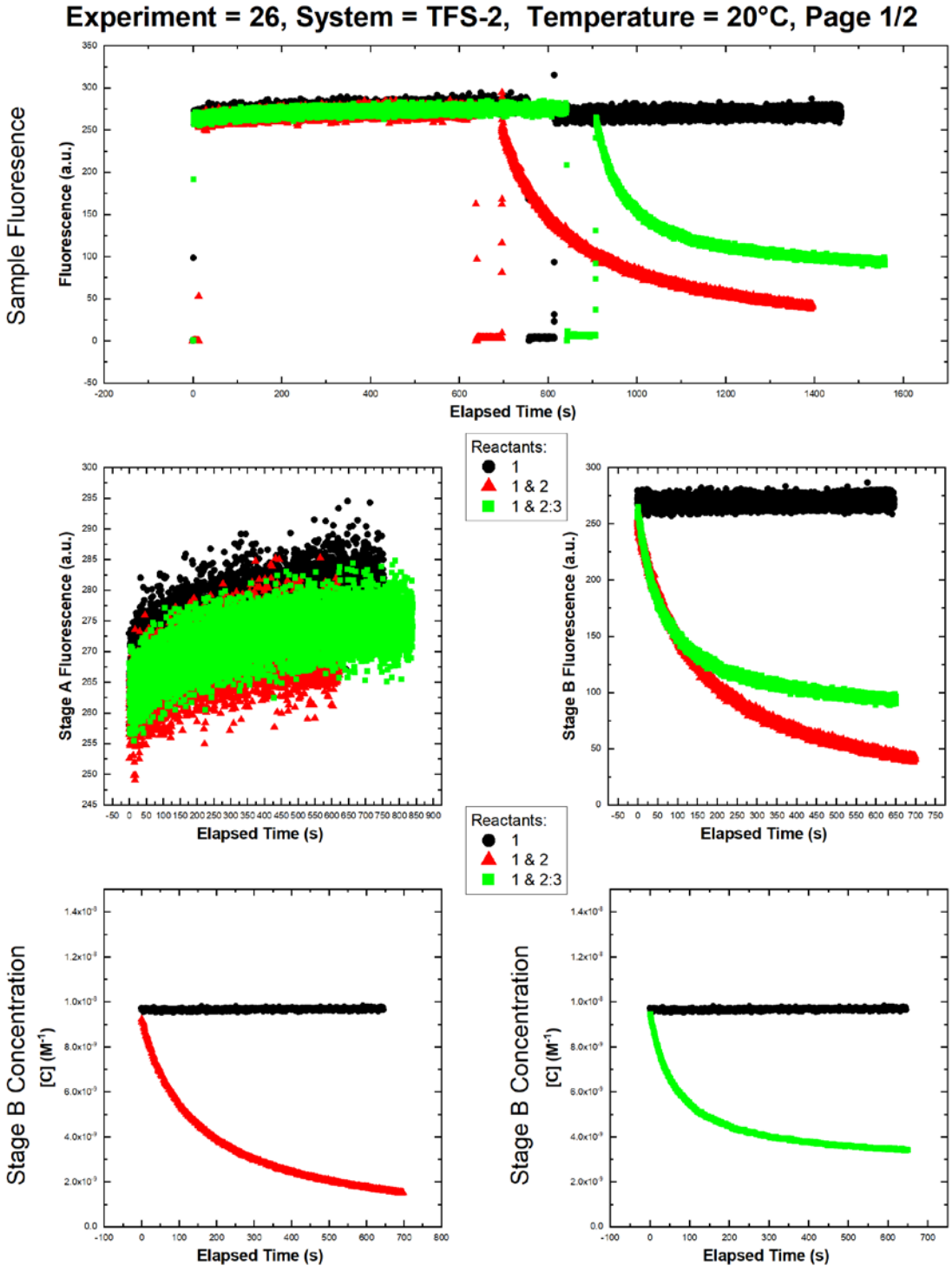

Experiment = 23, System = TFS-1, Temperature = 50°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

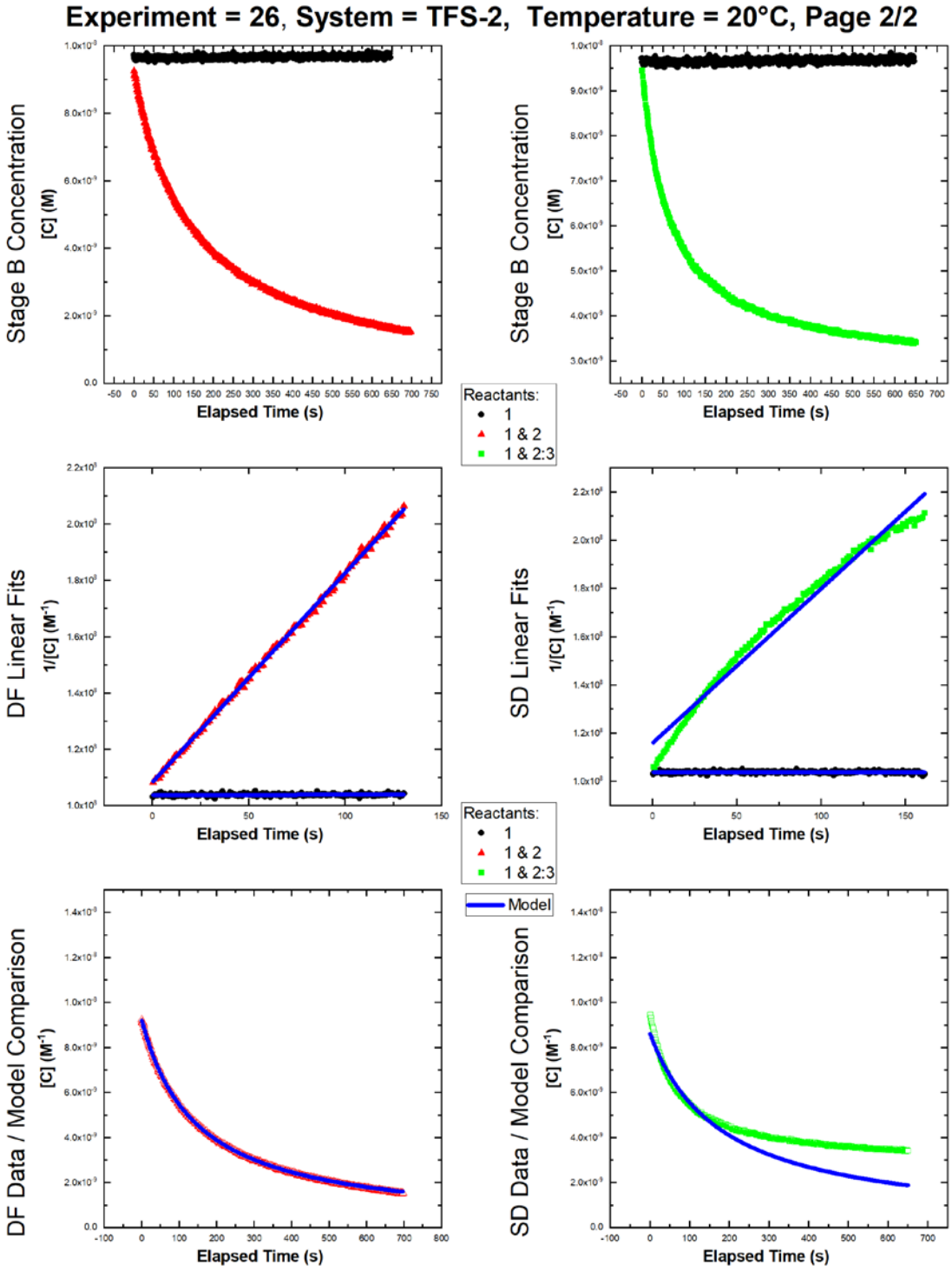

Experiment = 24, System = TFS-1, Temperature = 60°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

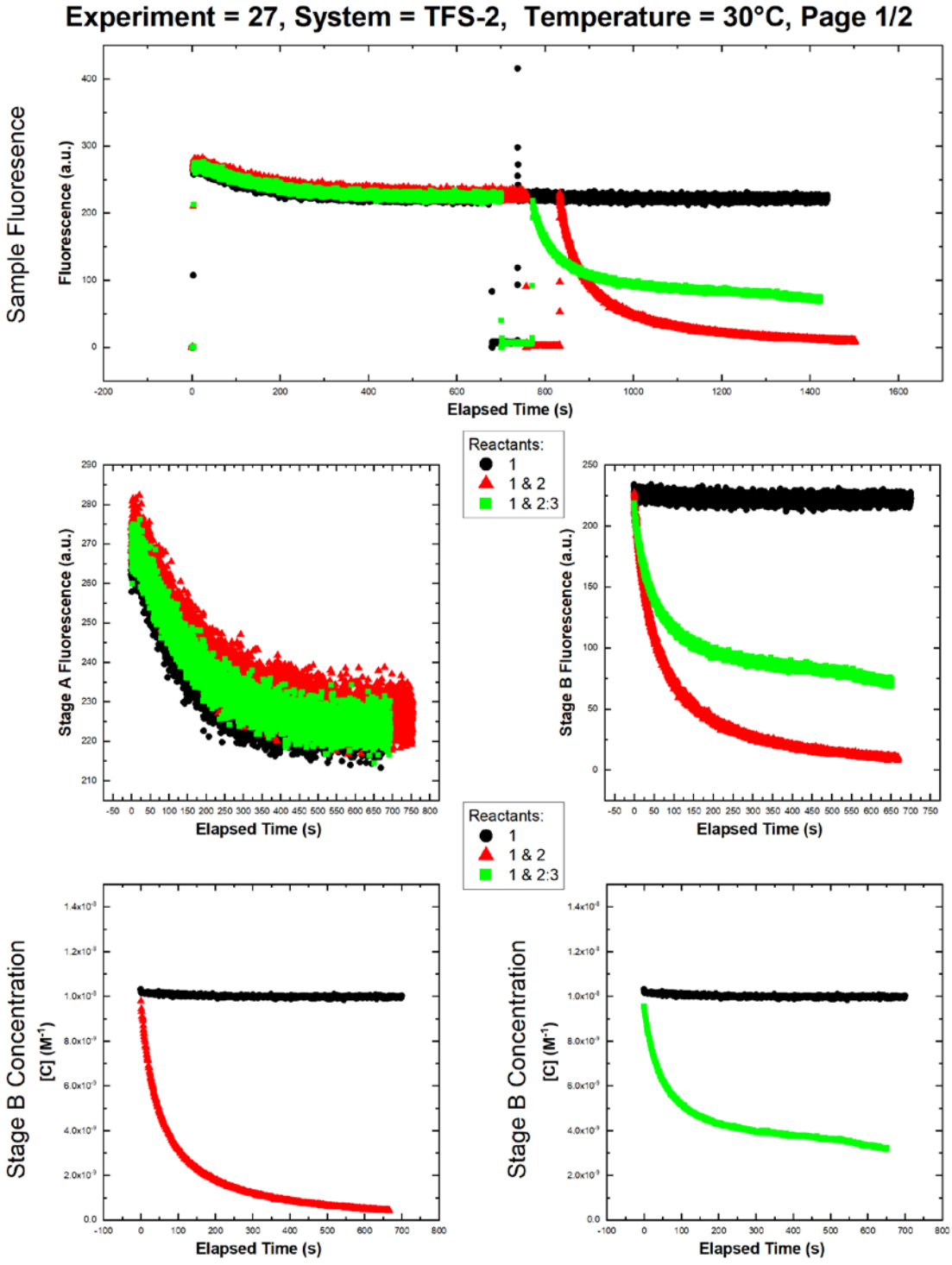

Experiment = 24, System = TFS-1, Temperature = 60°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

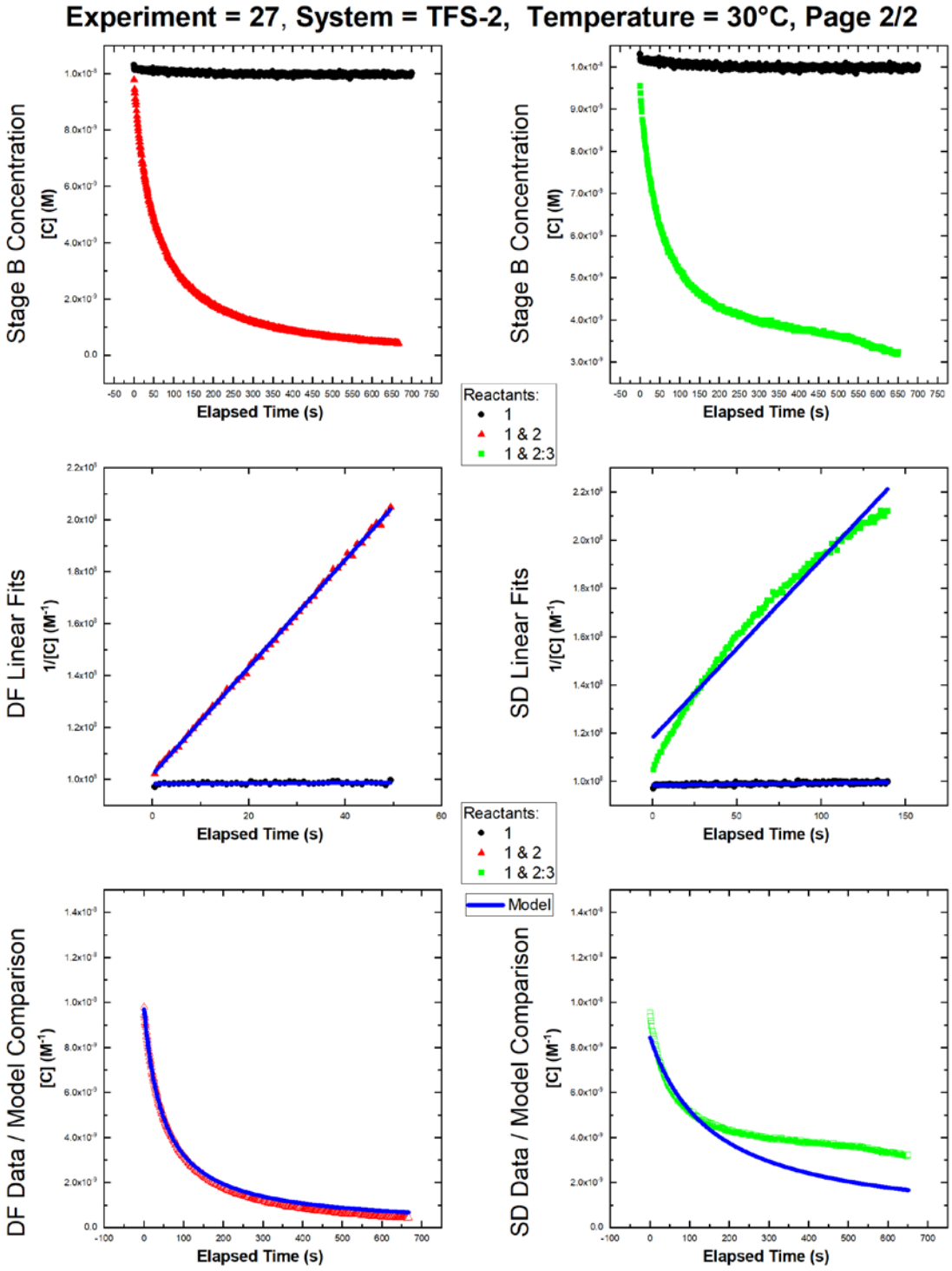

Experiment = 25, System = TFS-2, Temperature = 10°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

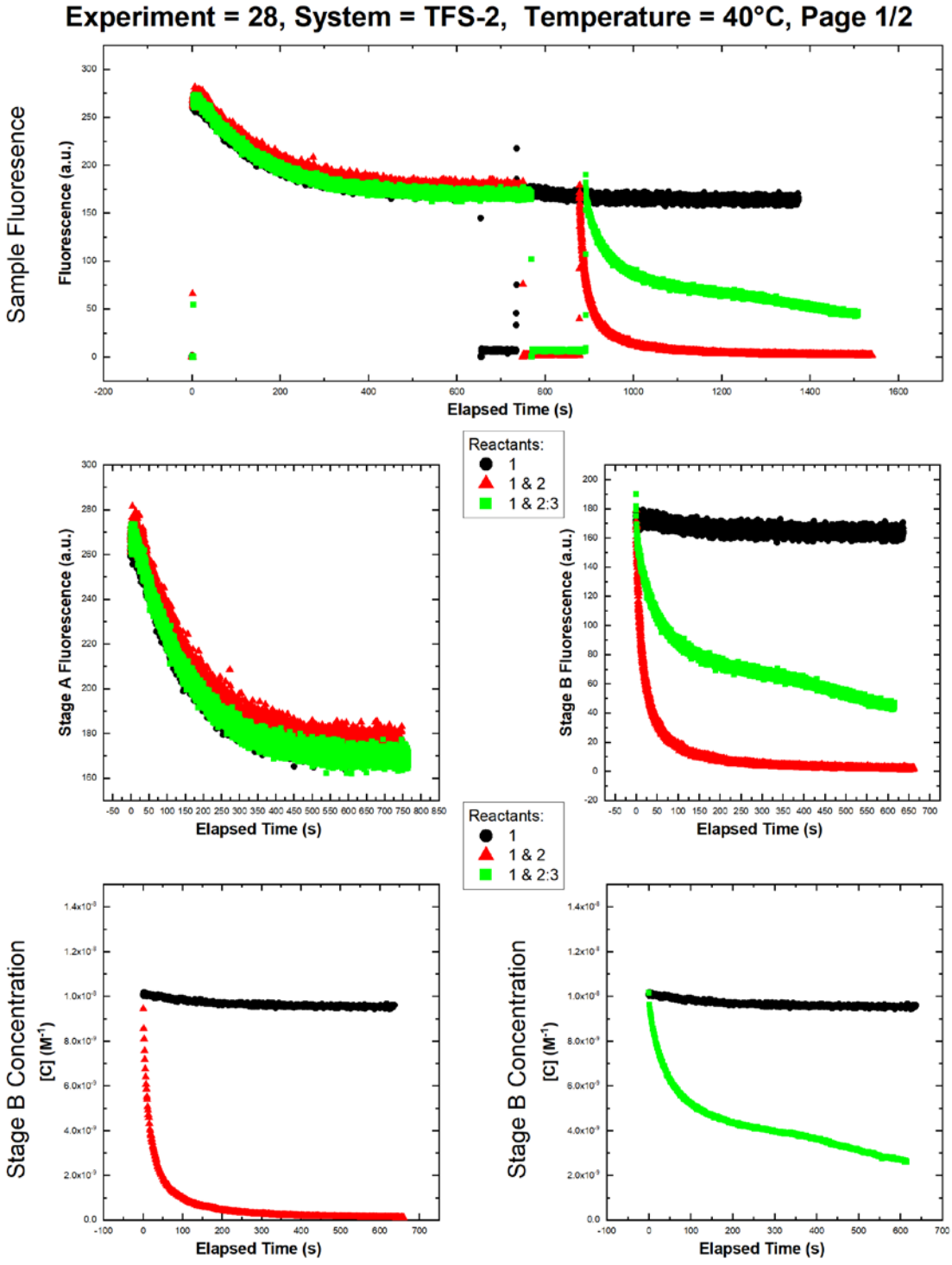

Experiment = 25, System = TFS-2, Temperature = 10°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

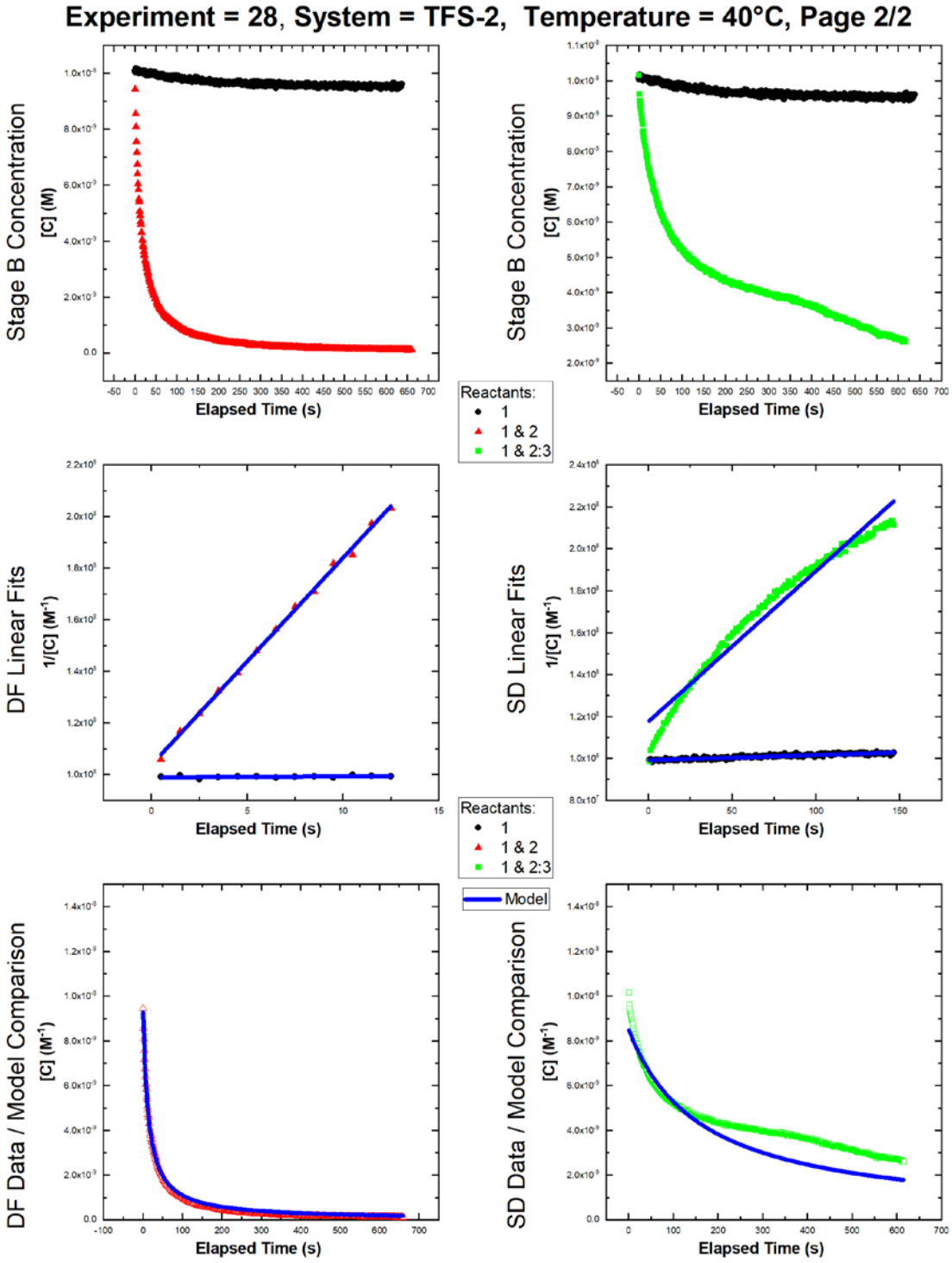

Experiment = 26, System = TFS-2, Temperature = 20°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

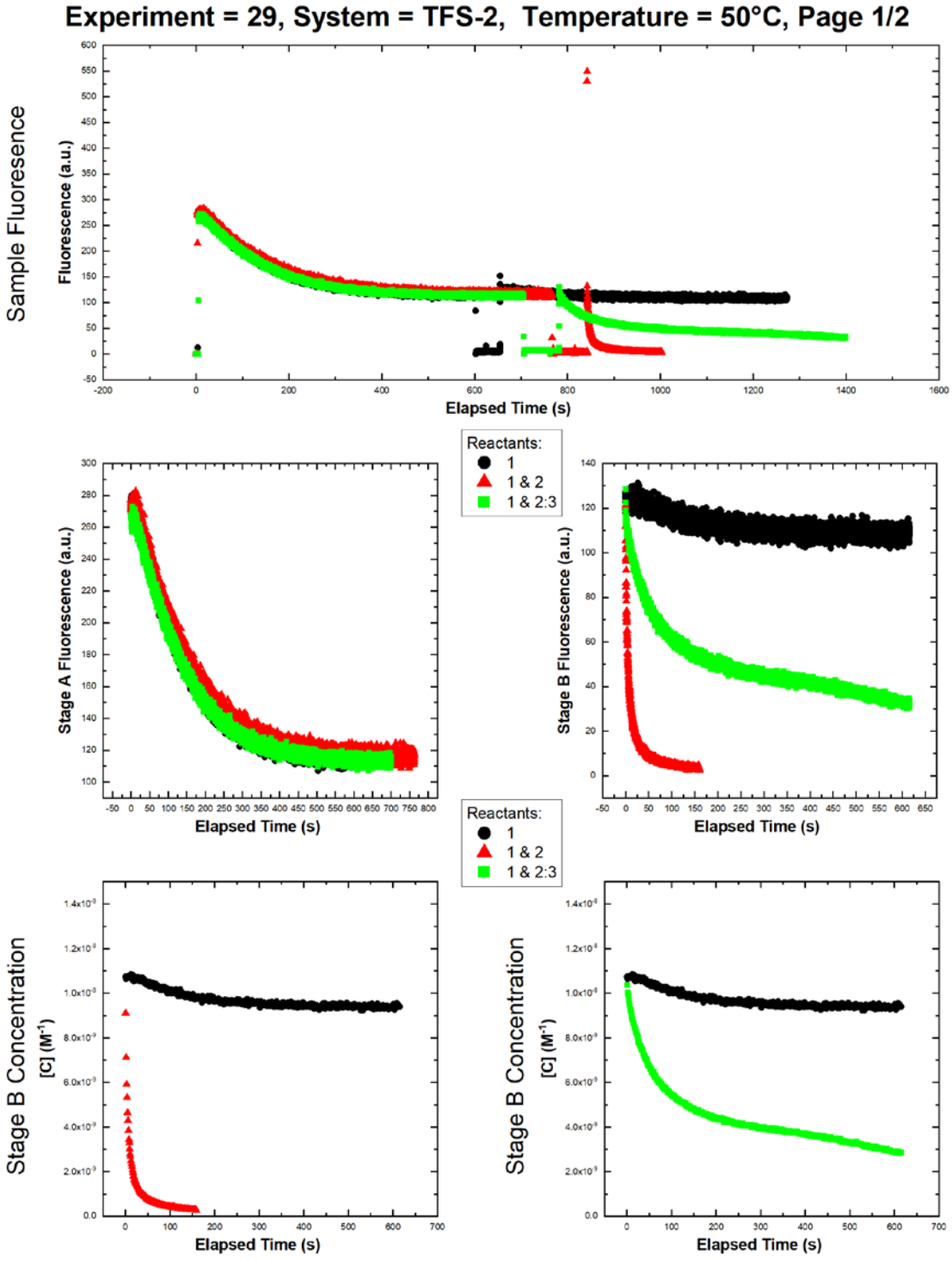

Experiment = 26, System = TFS-2, Temperature = 20°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

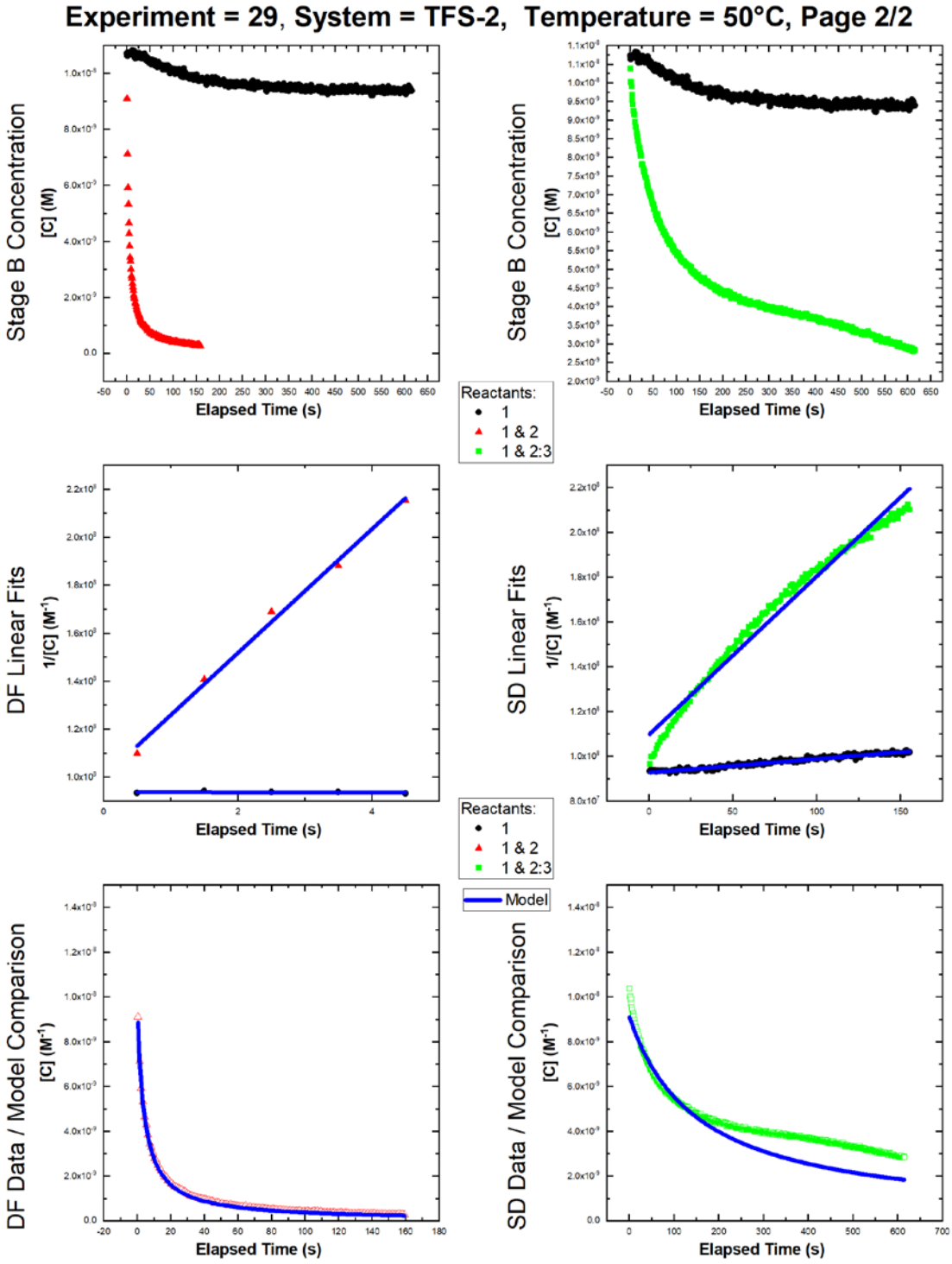

Experiment = 27, System = TFS-2, Temperature = 30°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

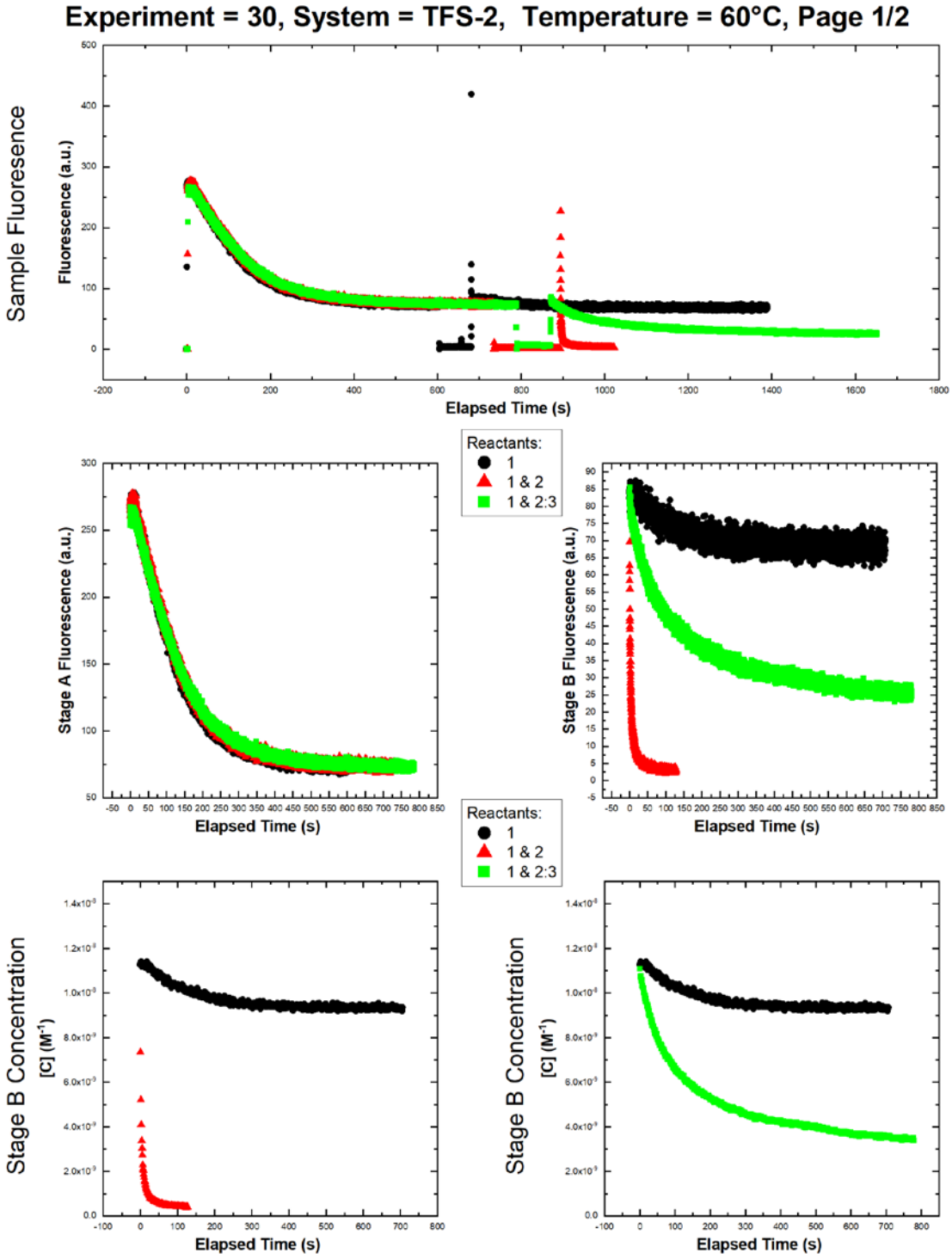

Experiment = 27, System = TFS-2, Temperature = 30°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

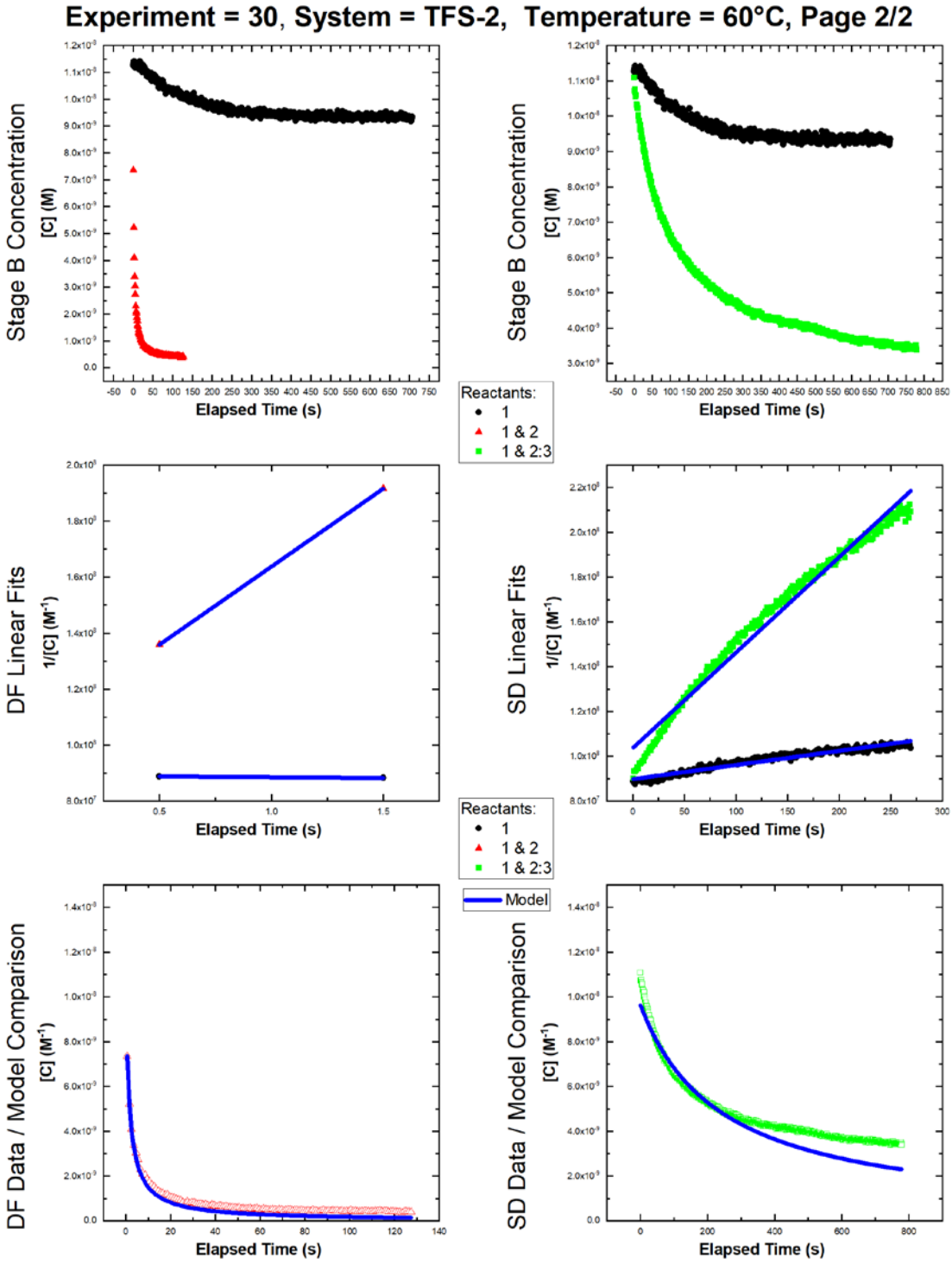

Experiment = 28, System = TFS-2, Temperature = 40°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

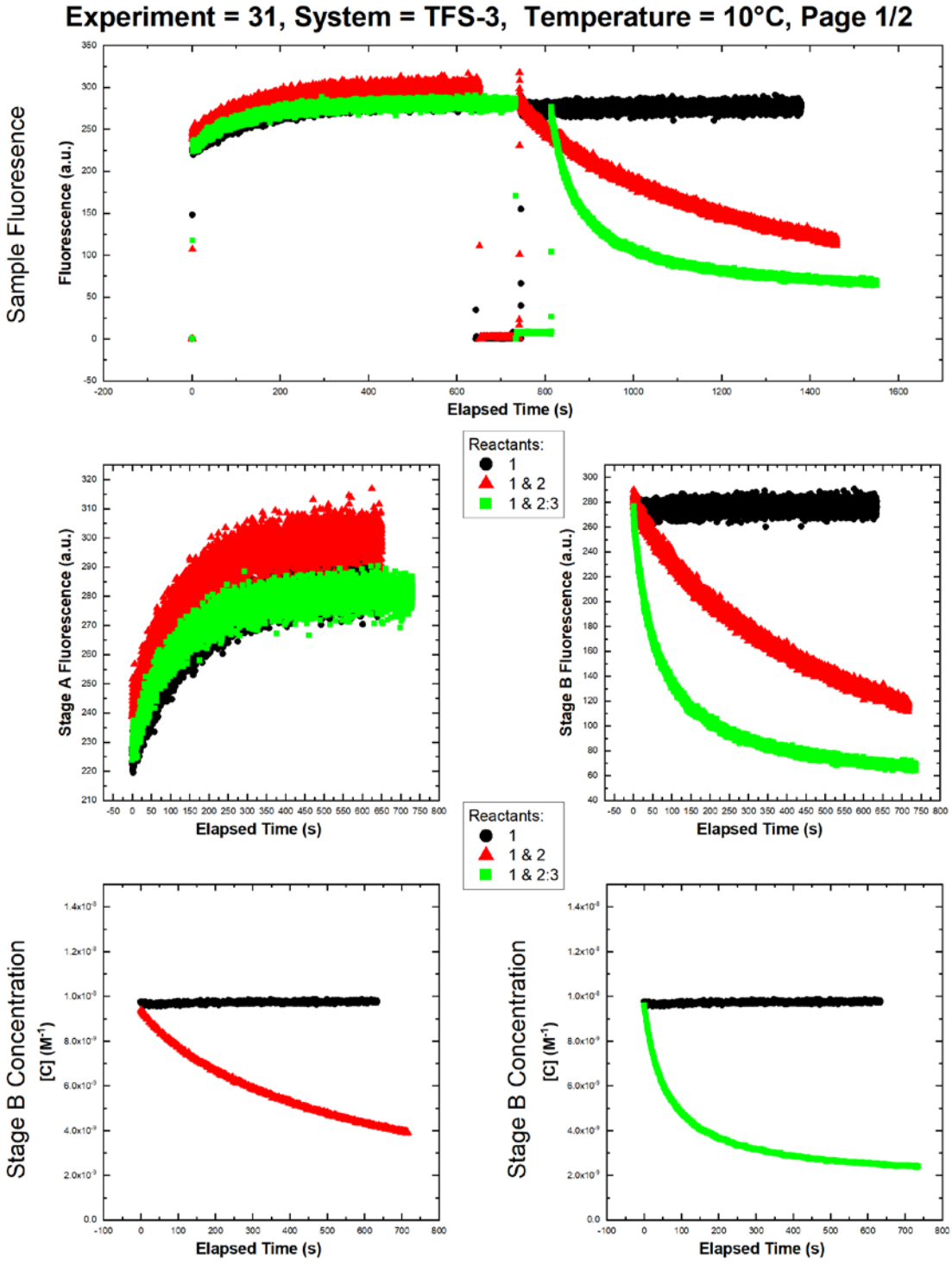

Experiment = 28, System = TFS-2, Temperature = 40°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

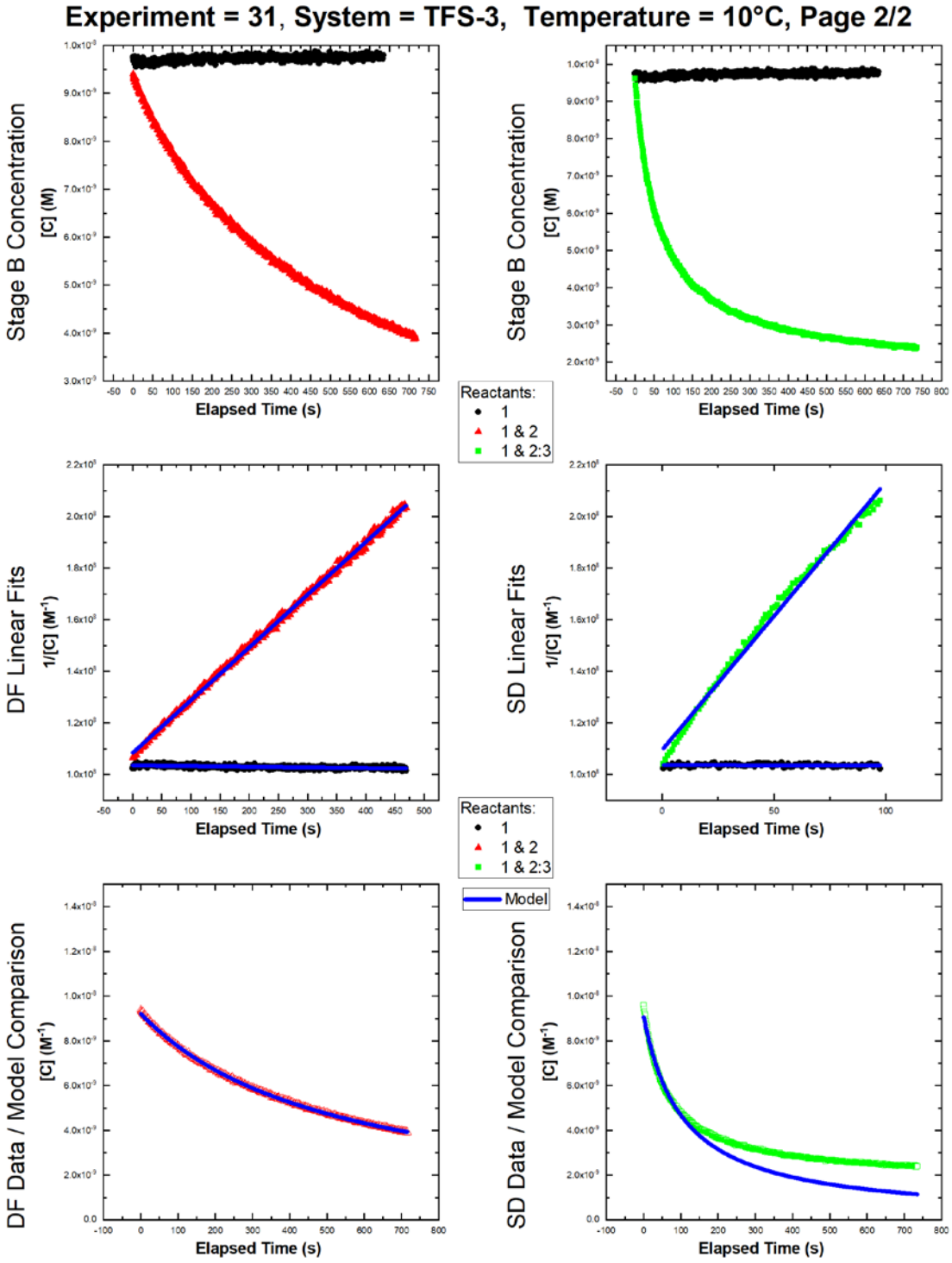

Experiment = 29, System = TFS-2, Temperature = 50°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

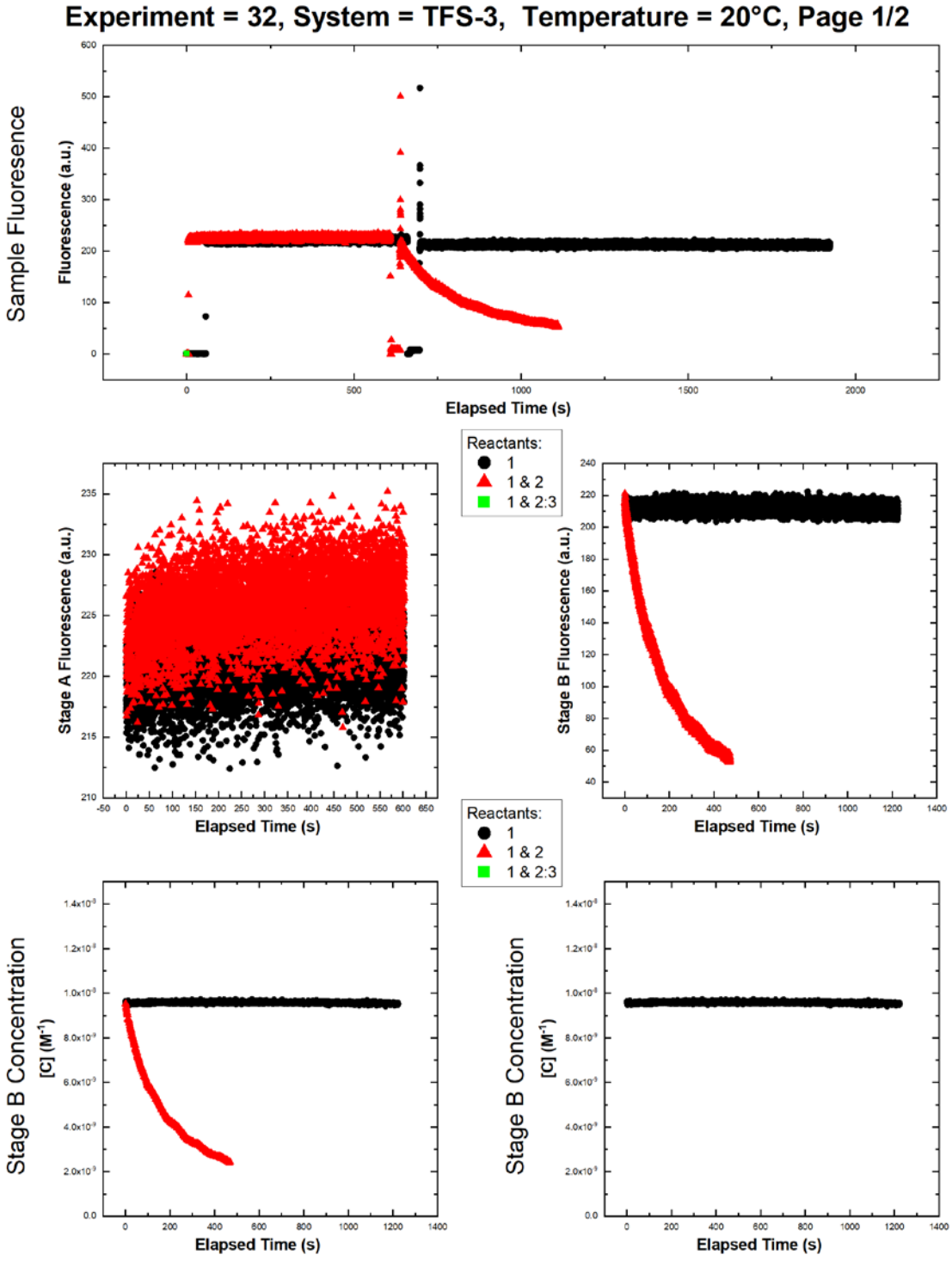

Experiment = 29, System = TFS-2, Temperature = 50°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

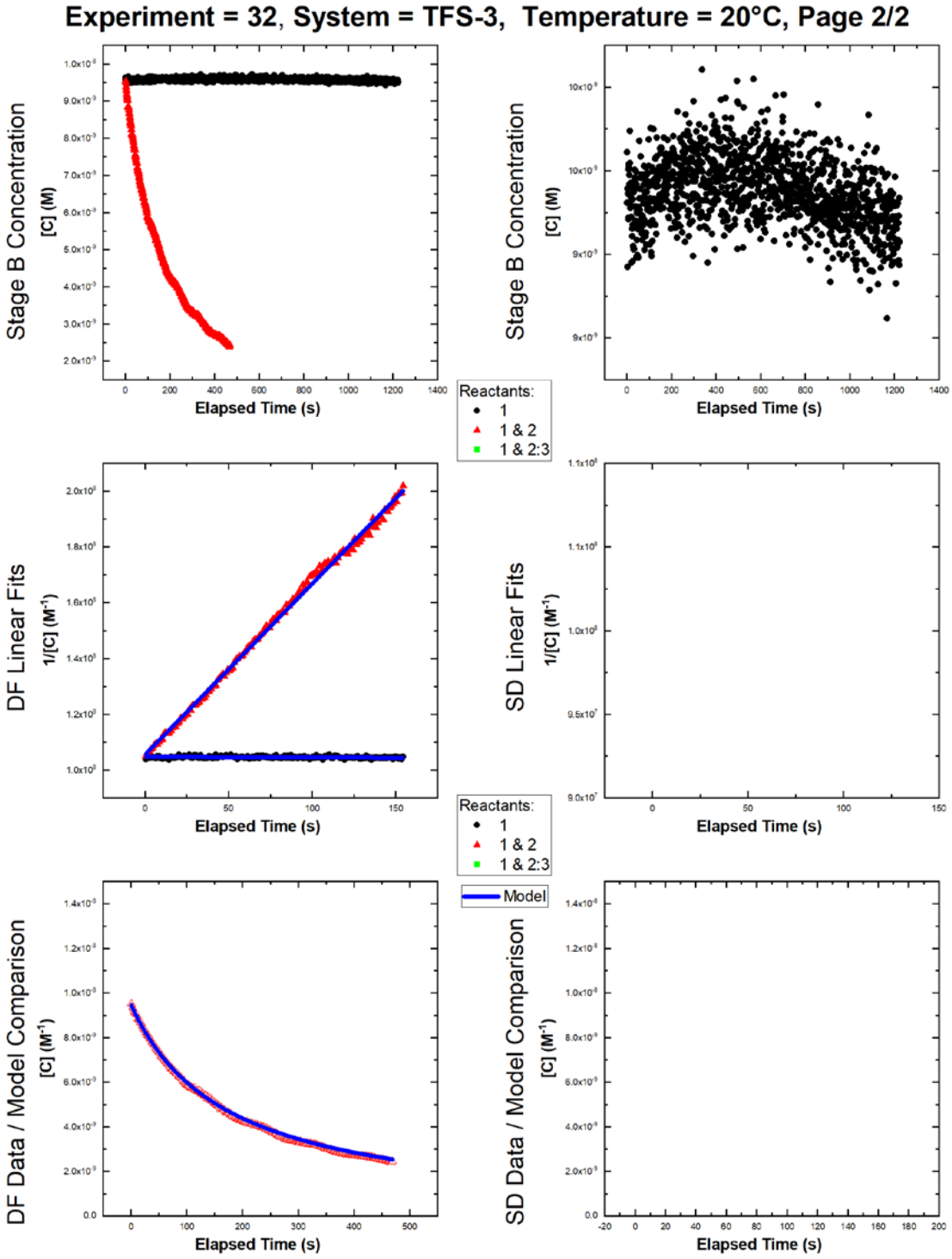

Experiment = 30, System = TFS-2, Temperature = 60°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

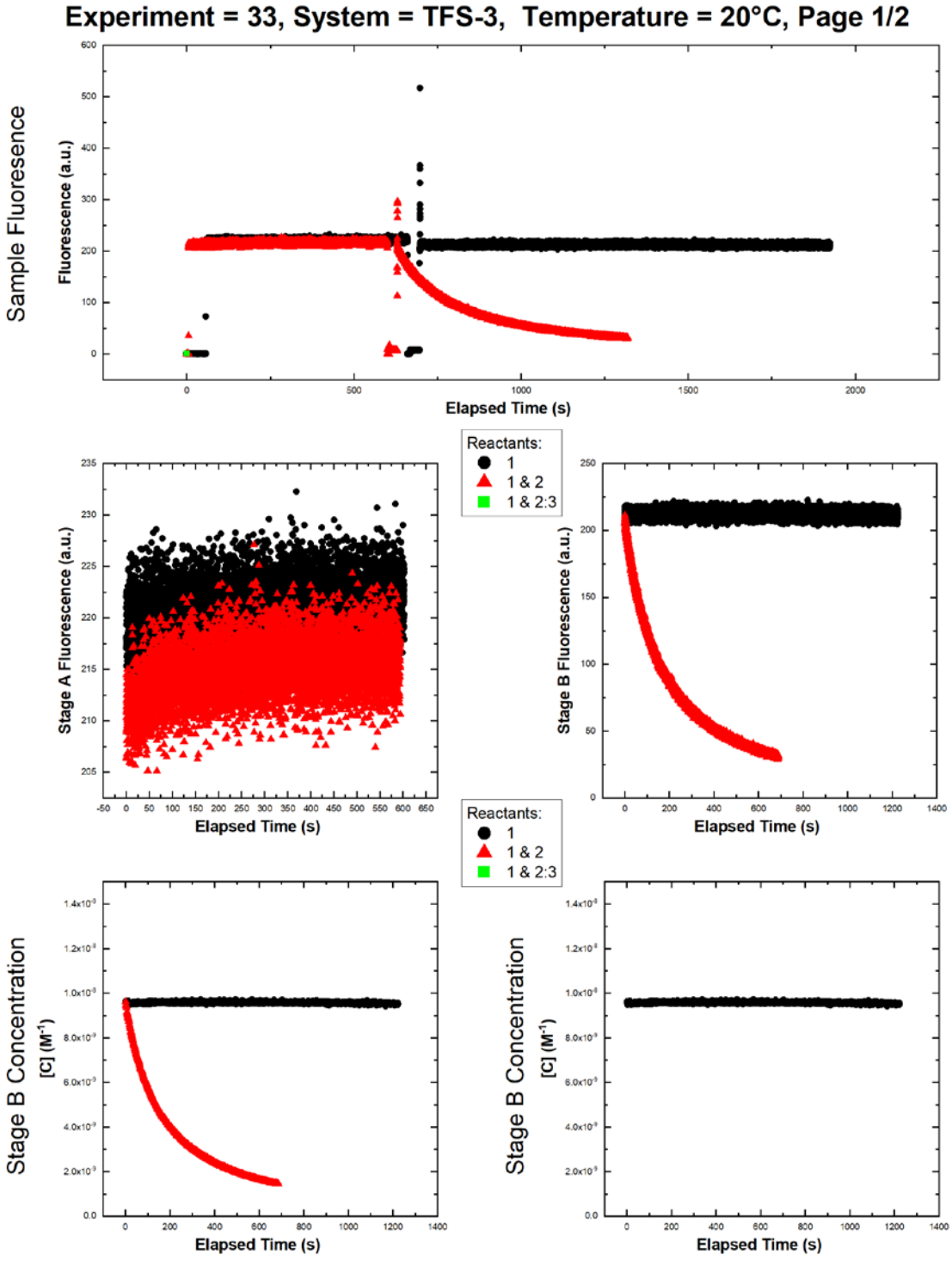

Experiment = 30, System = TFS-2, Temperature = 60°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

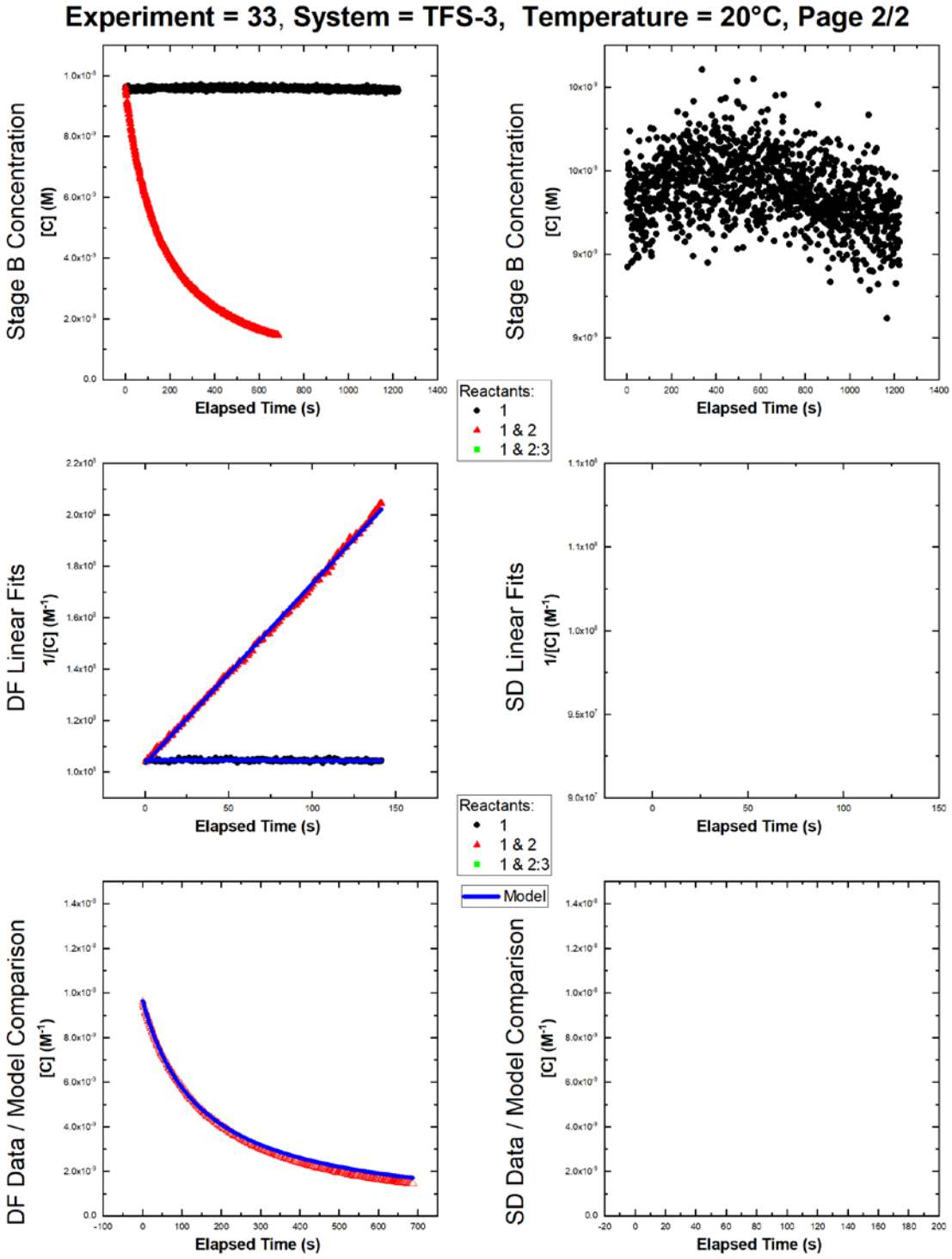

Experiment = 31, System = TFS-3, Temperature = 10°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

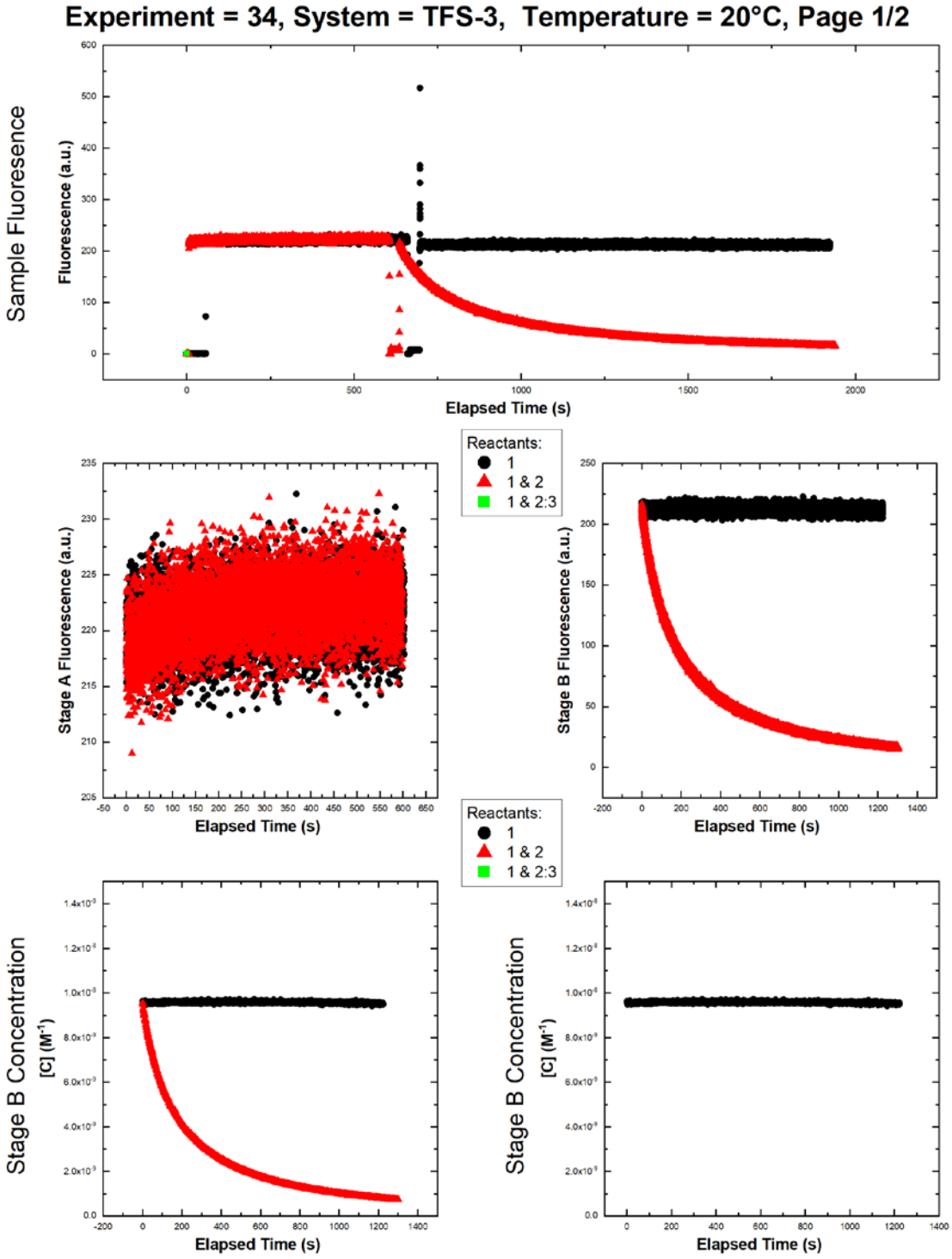

Experiment = 31, System = TFS-3, Temperature = 10°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

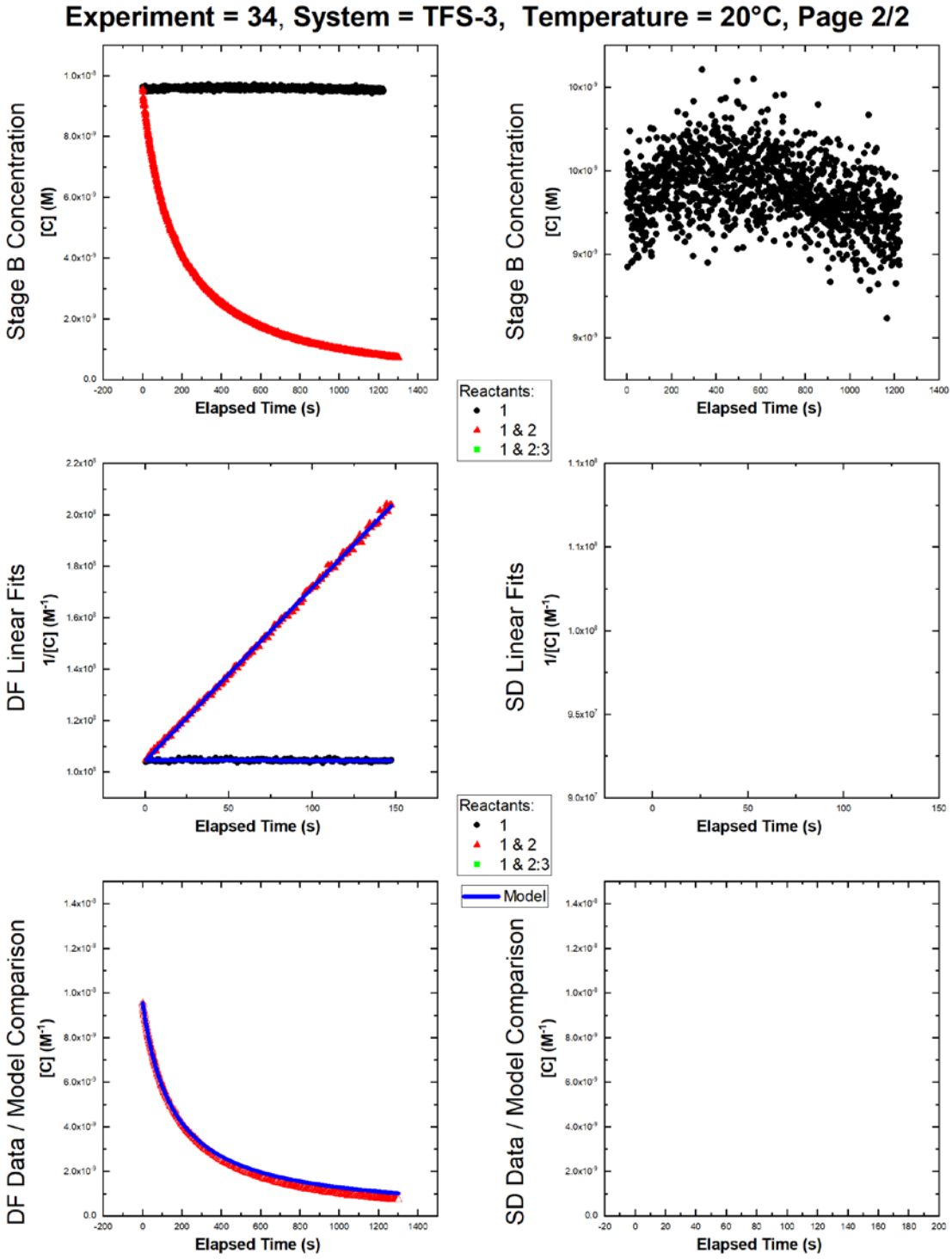

Experiment = 32, System = TFS-3, Temperature = 20°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

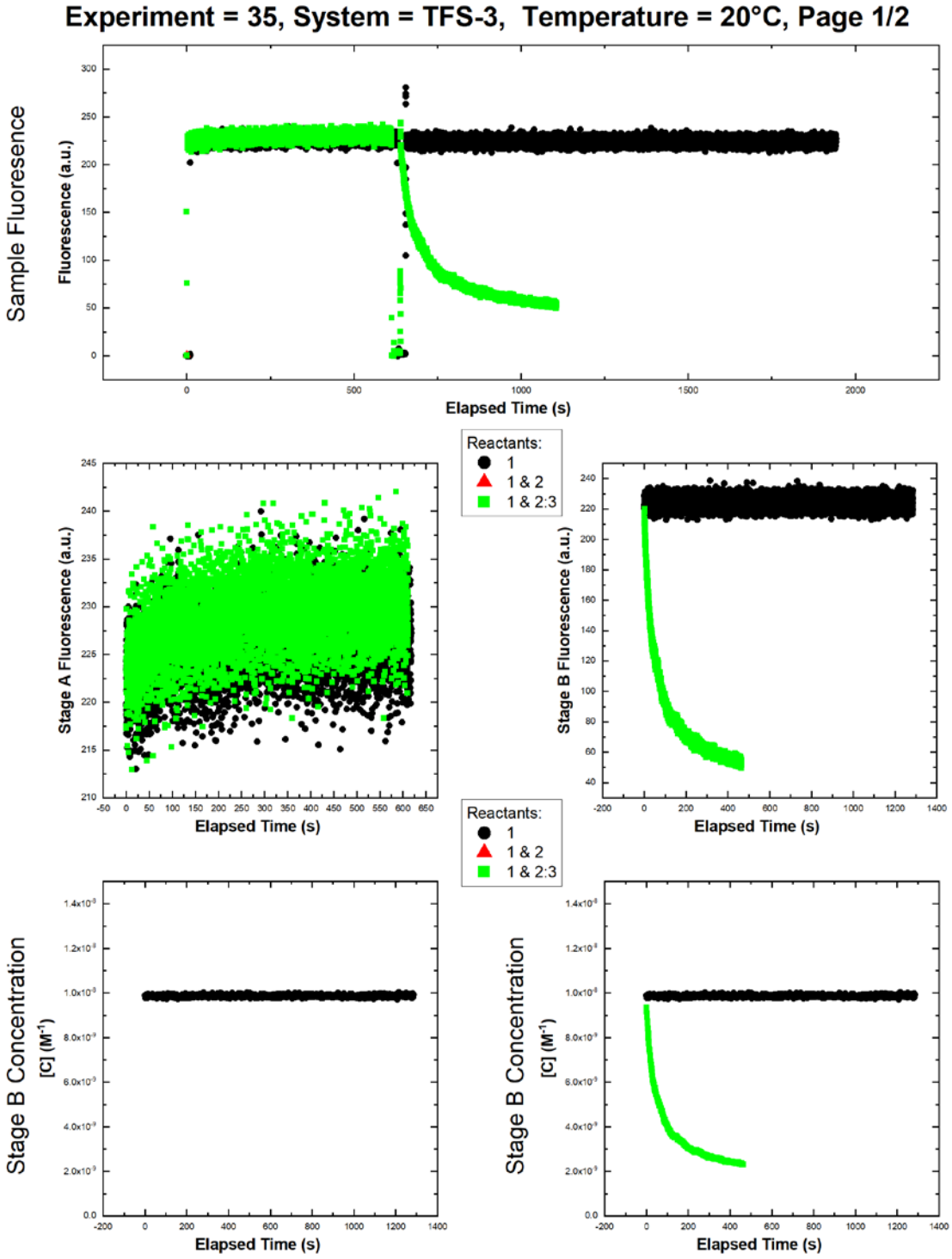

Experiment = 32, System = TFS-3, Temperature = 20°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

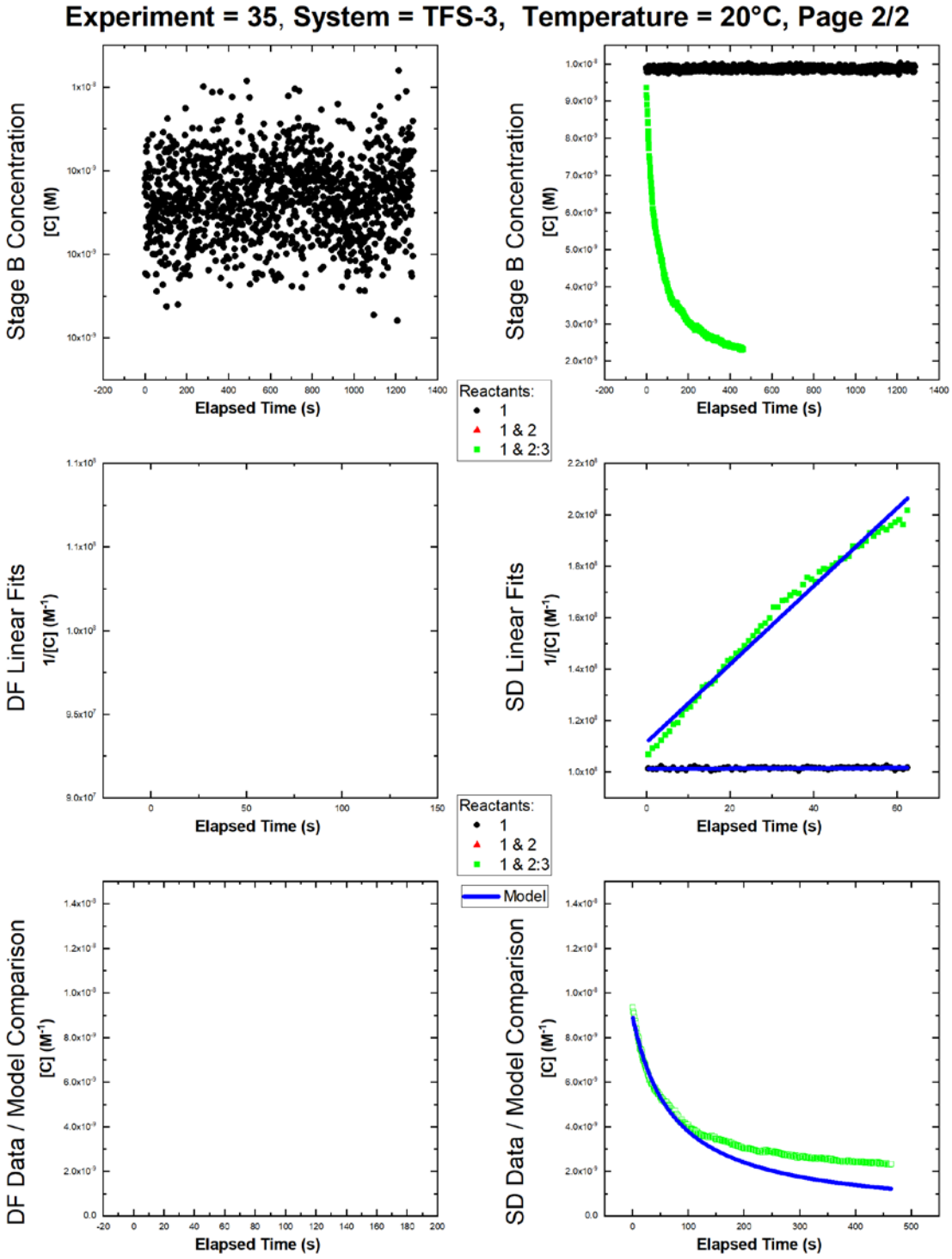

Experiment = 33, System = TFS-3, Temperature = 20°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

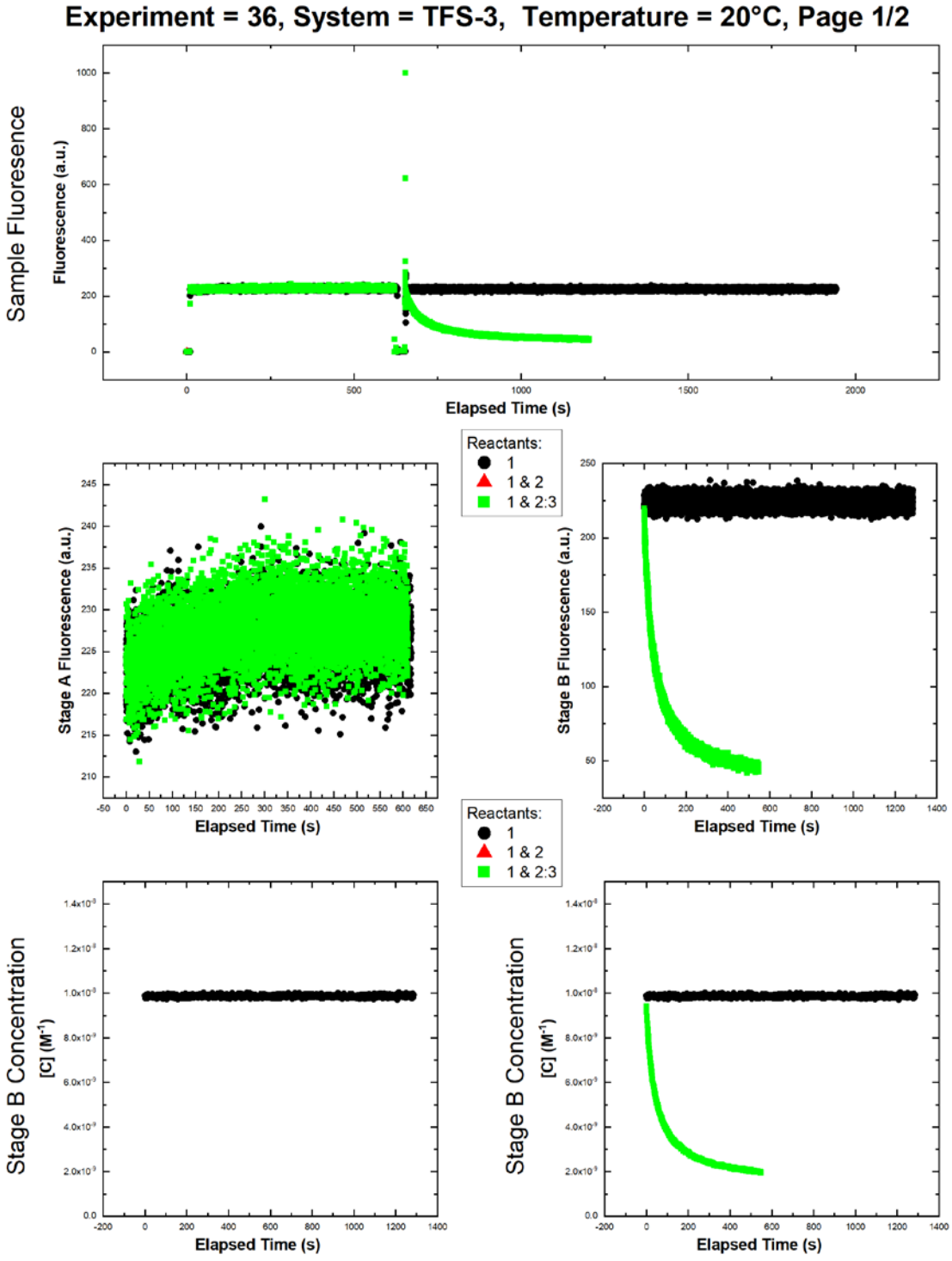

Experiment = 33, System = TFS-3, Temperature = 20°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

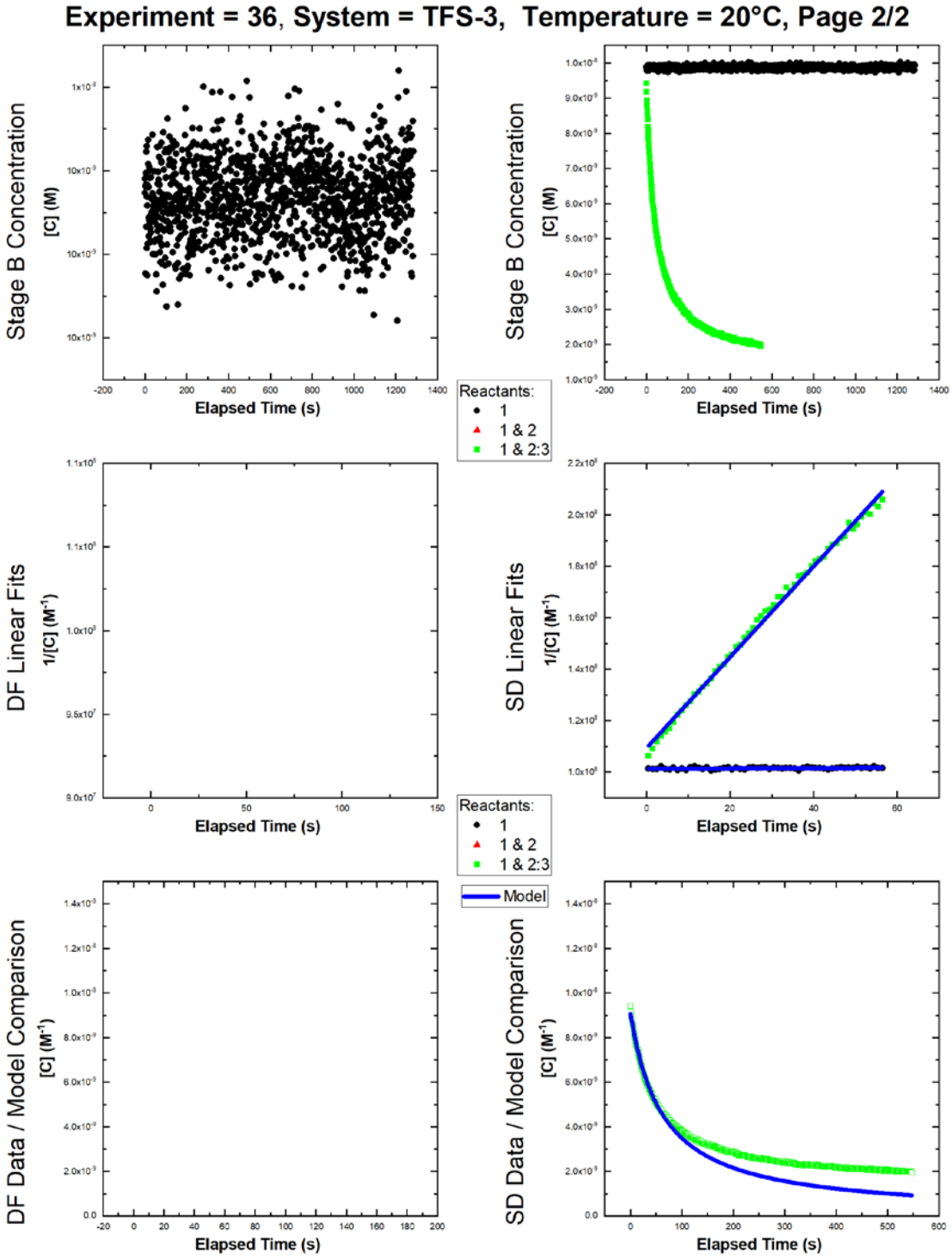

Experiment = 34, System = TFS-3, Temperature = 20°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

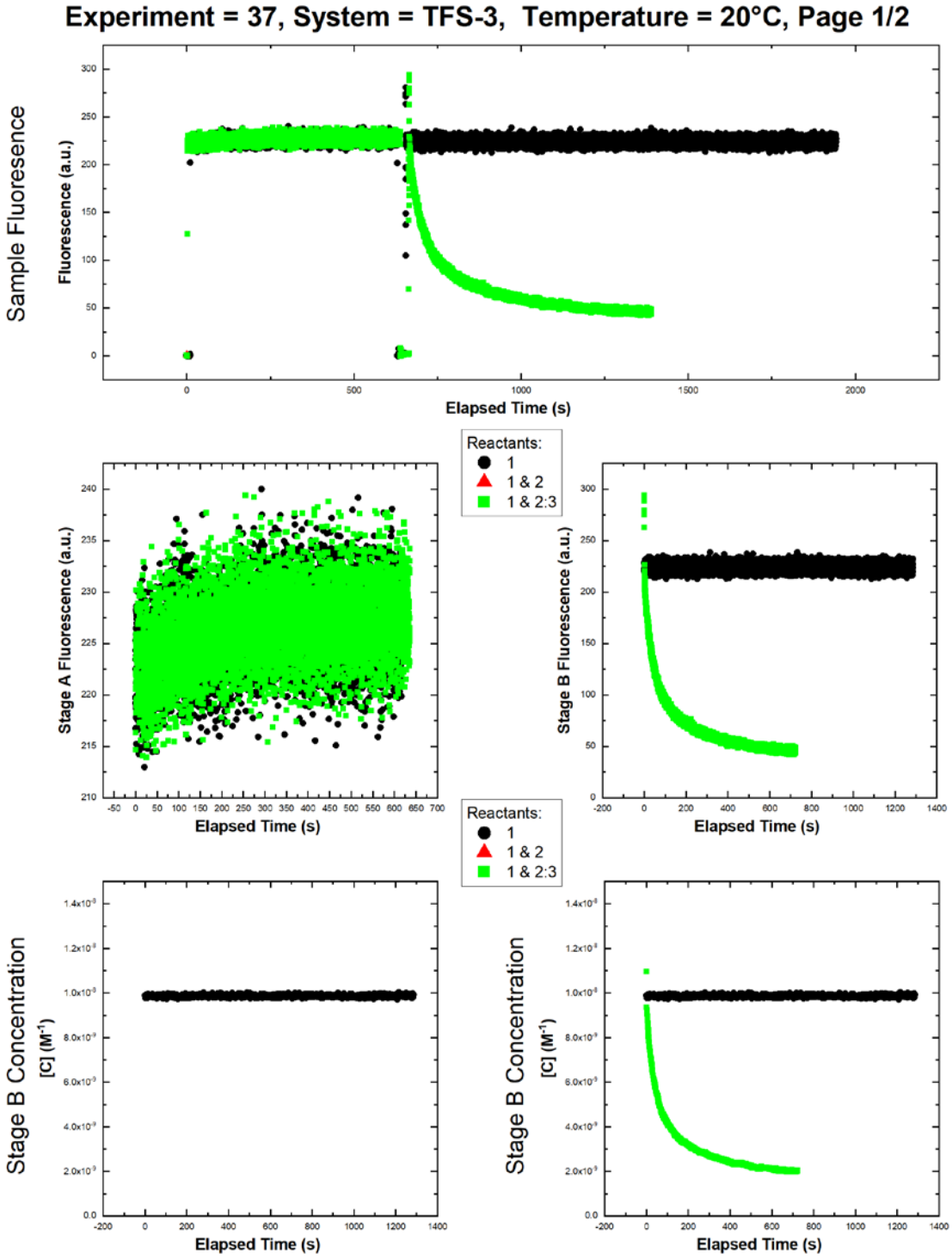

Experiment = 34, System = TFS-3, Temperature = 20°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

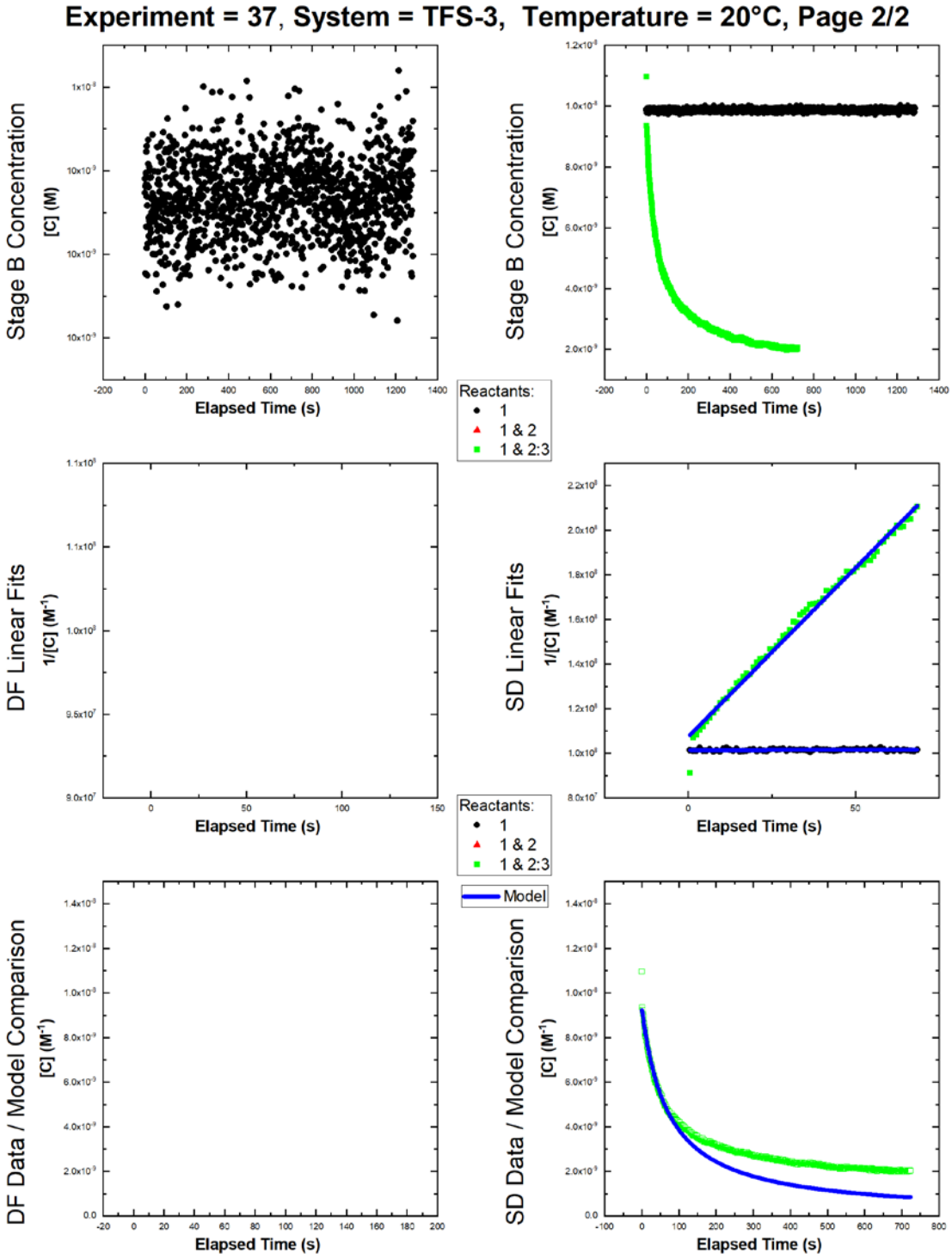

Experiment = 35, System = TFS-3, Temperature = 20°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

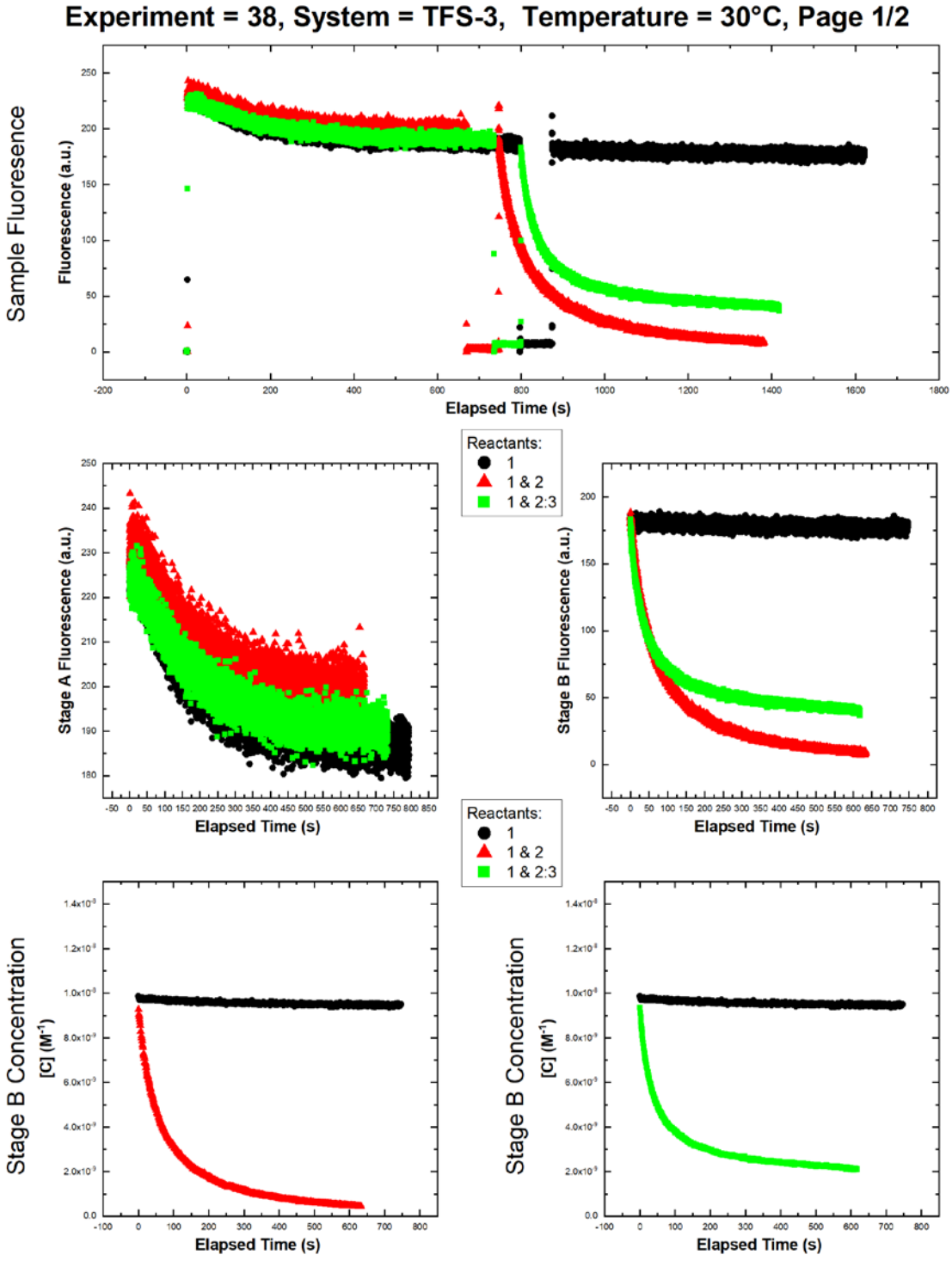

Experiment = 35, System = TFS-3, Temperature = 20°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

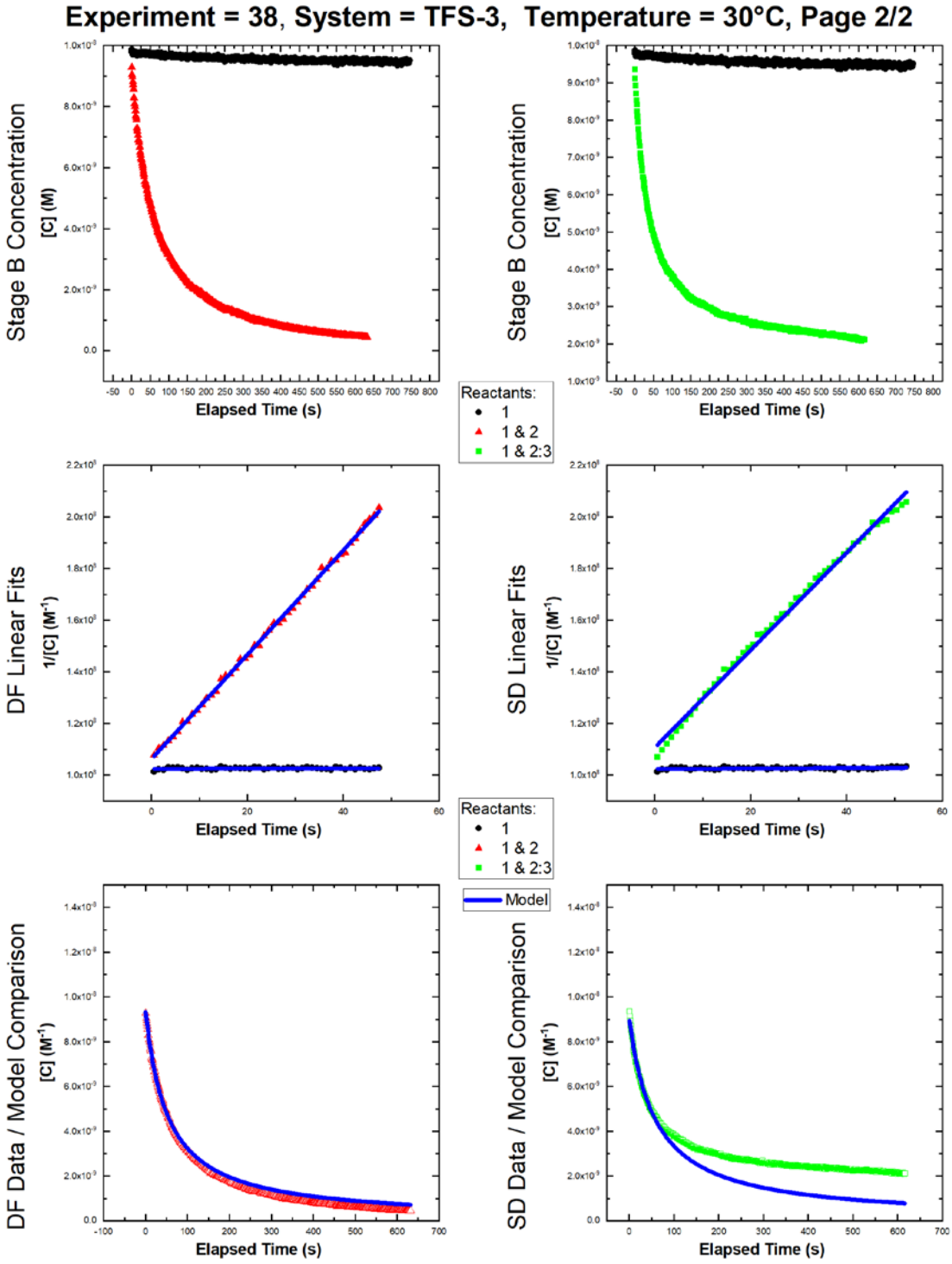

Experiment = 36, System = TFS-3, Temperature = 20°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

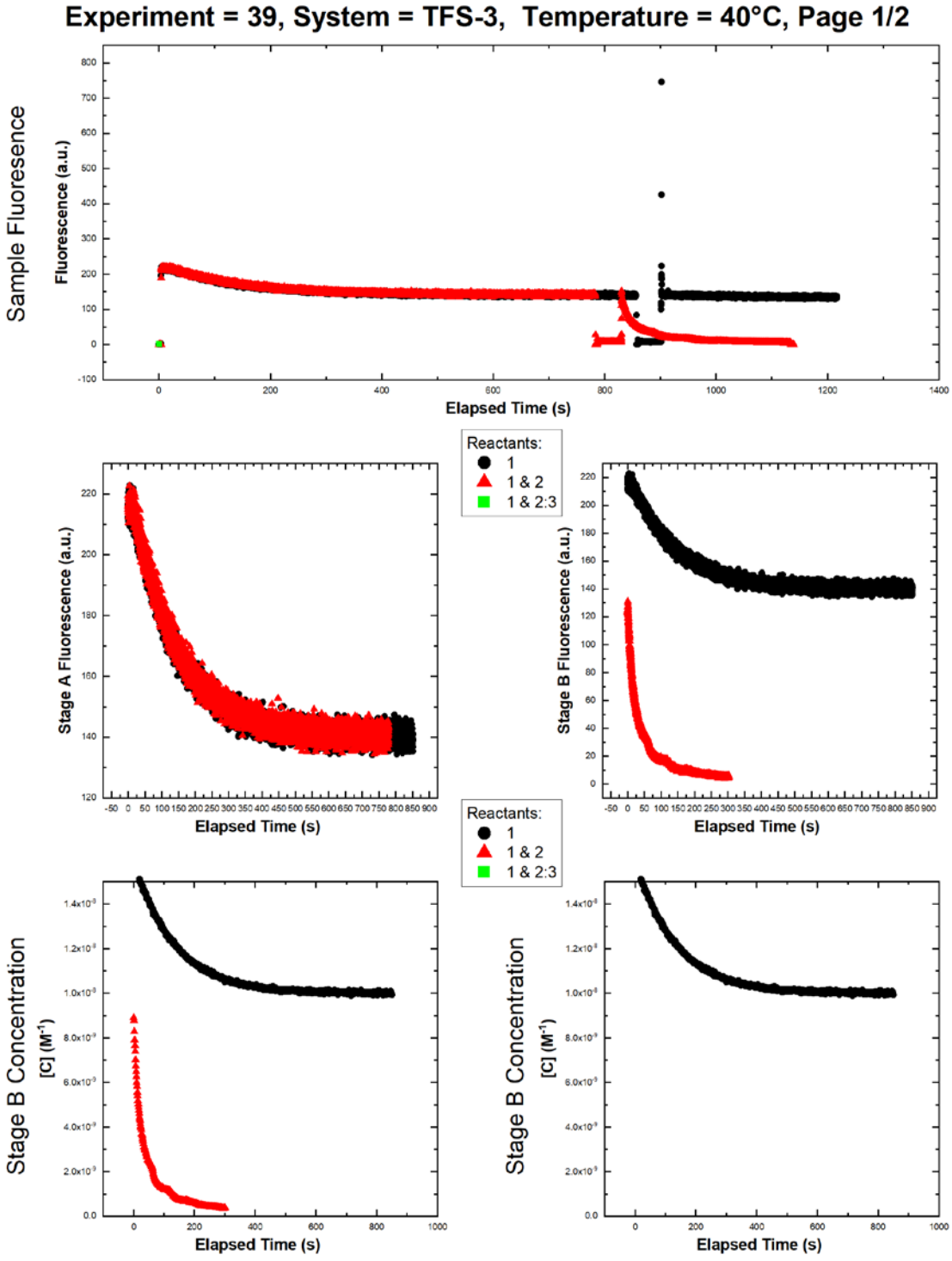

Experiment = 36, System = TFS-3, Temperature = 20°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

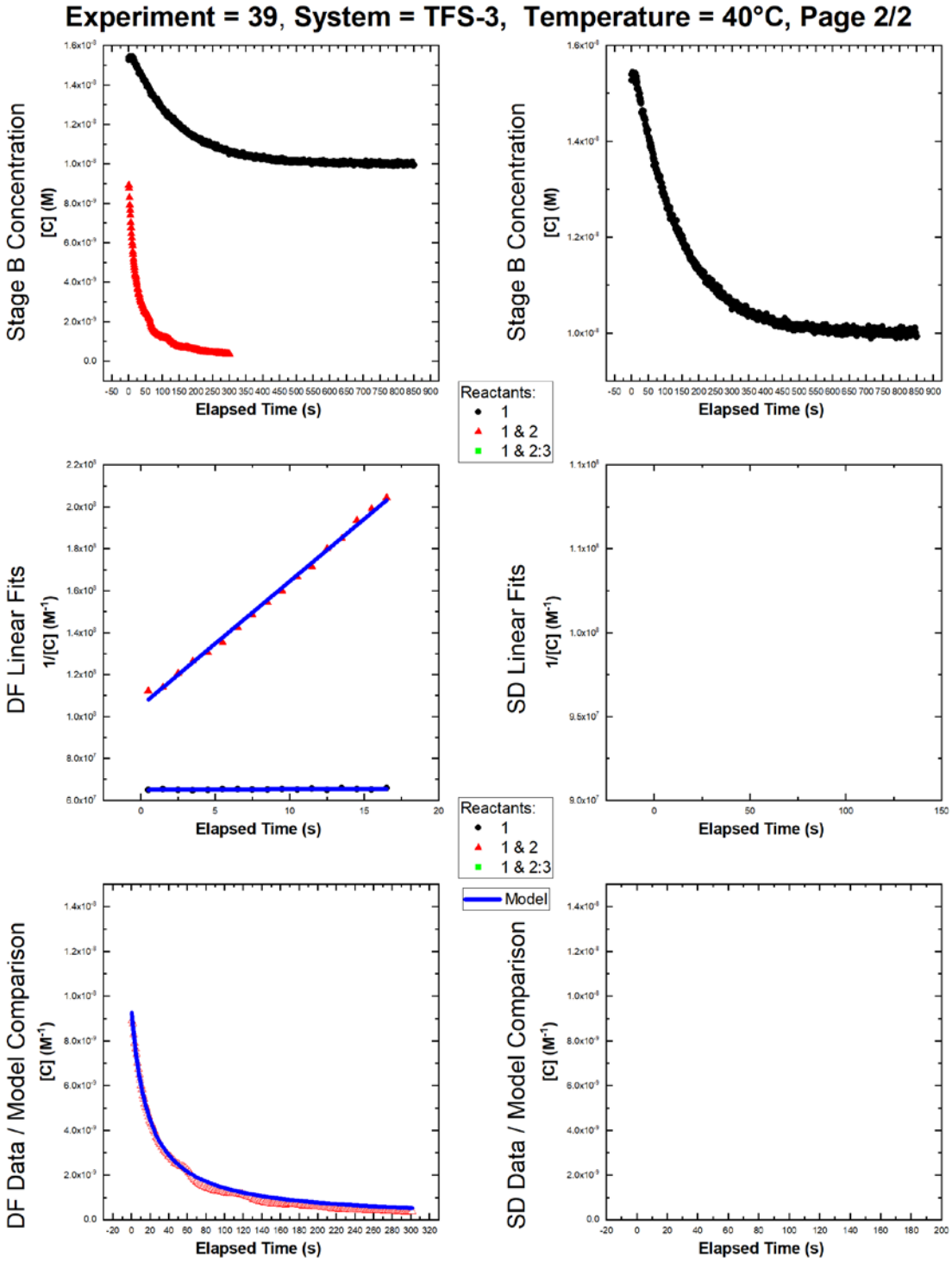

Experiment = 37, System = TFS-3, Temperature = 20°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

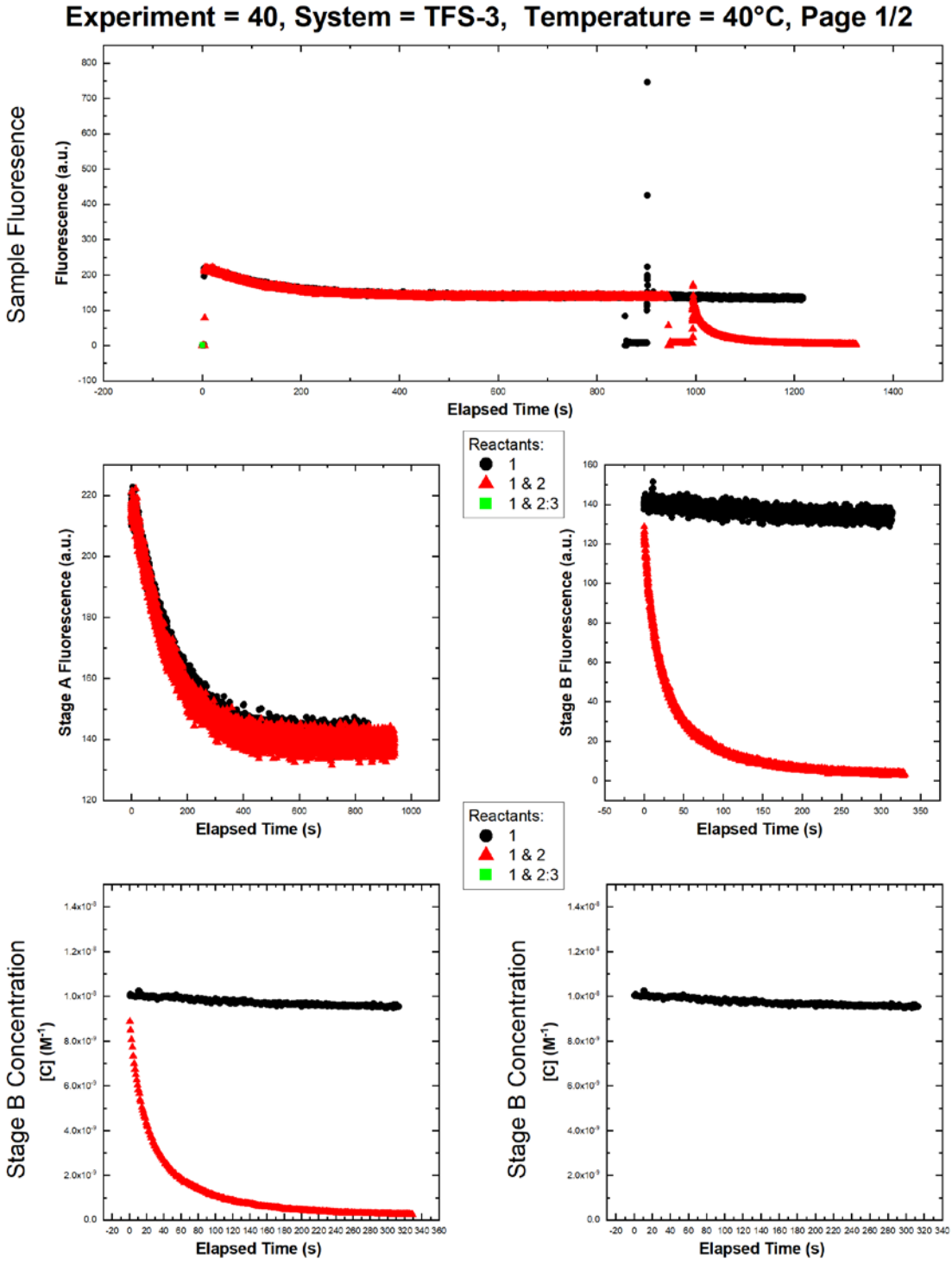

Experiment = 37, System = TFS-3, Temperature = 20°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

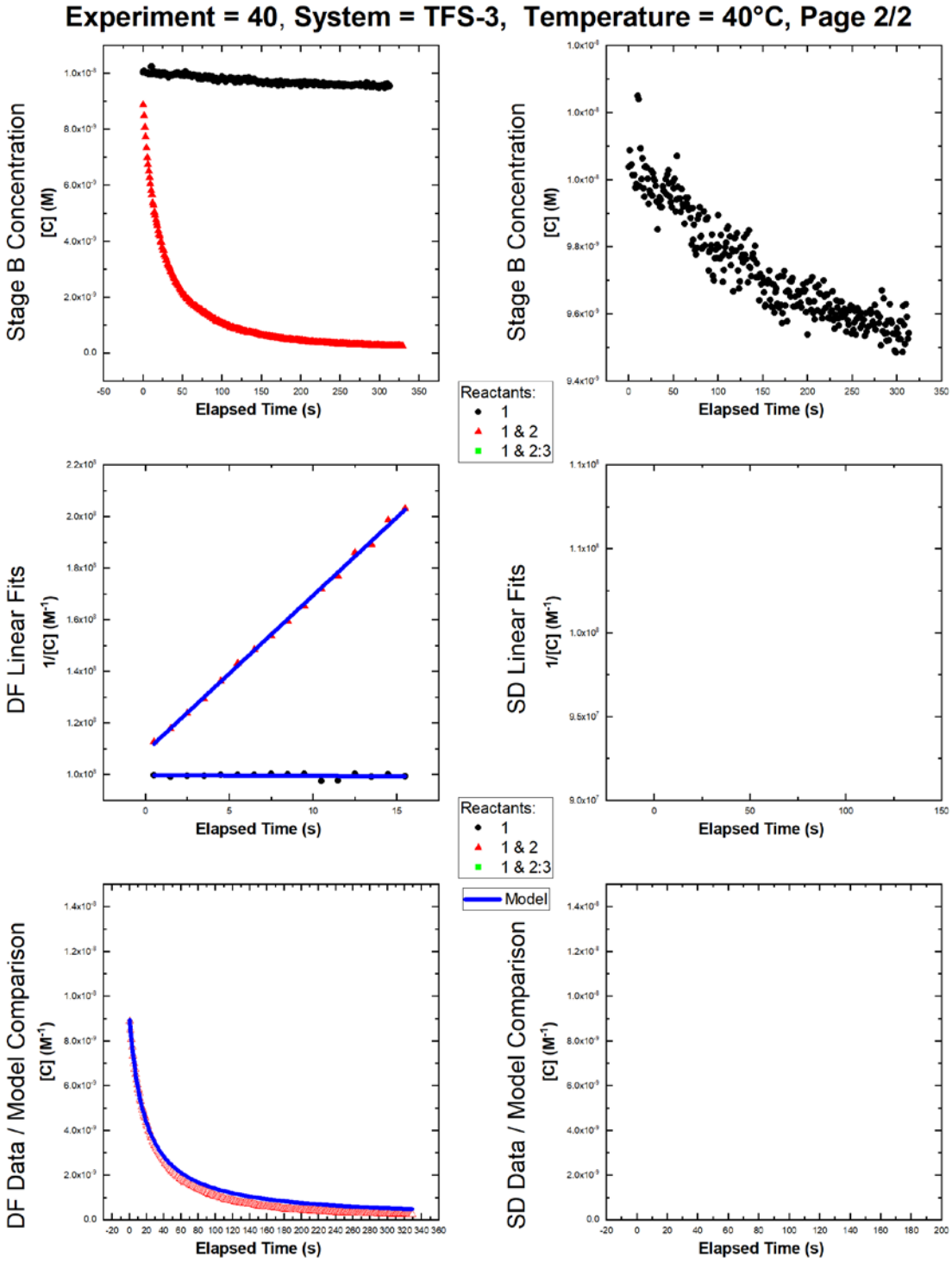

Experiment = 38, System = TFS-3, Temperature = 30°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

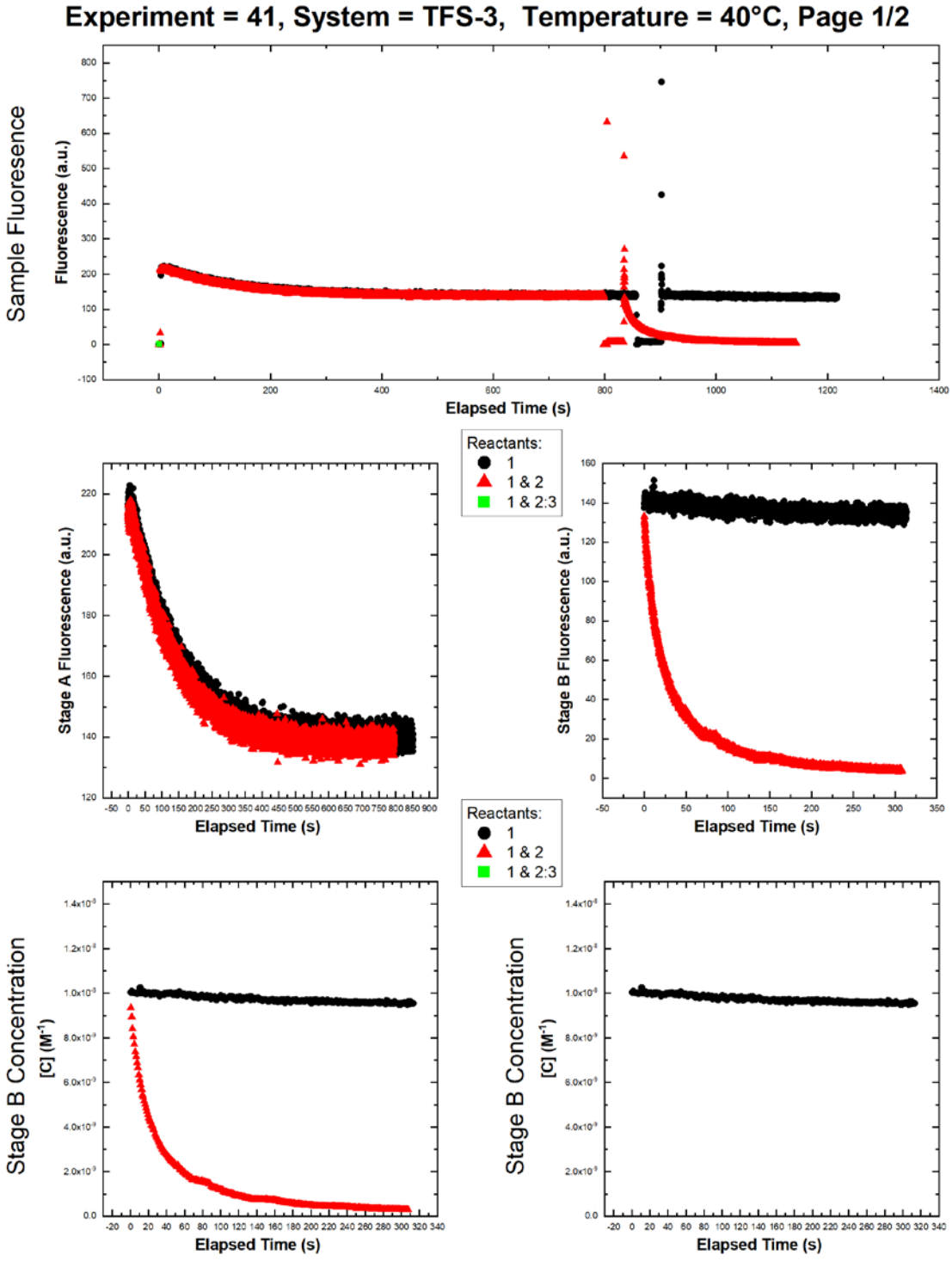

Experiment = 38, System = TFS-3, Temperature = 30°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

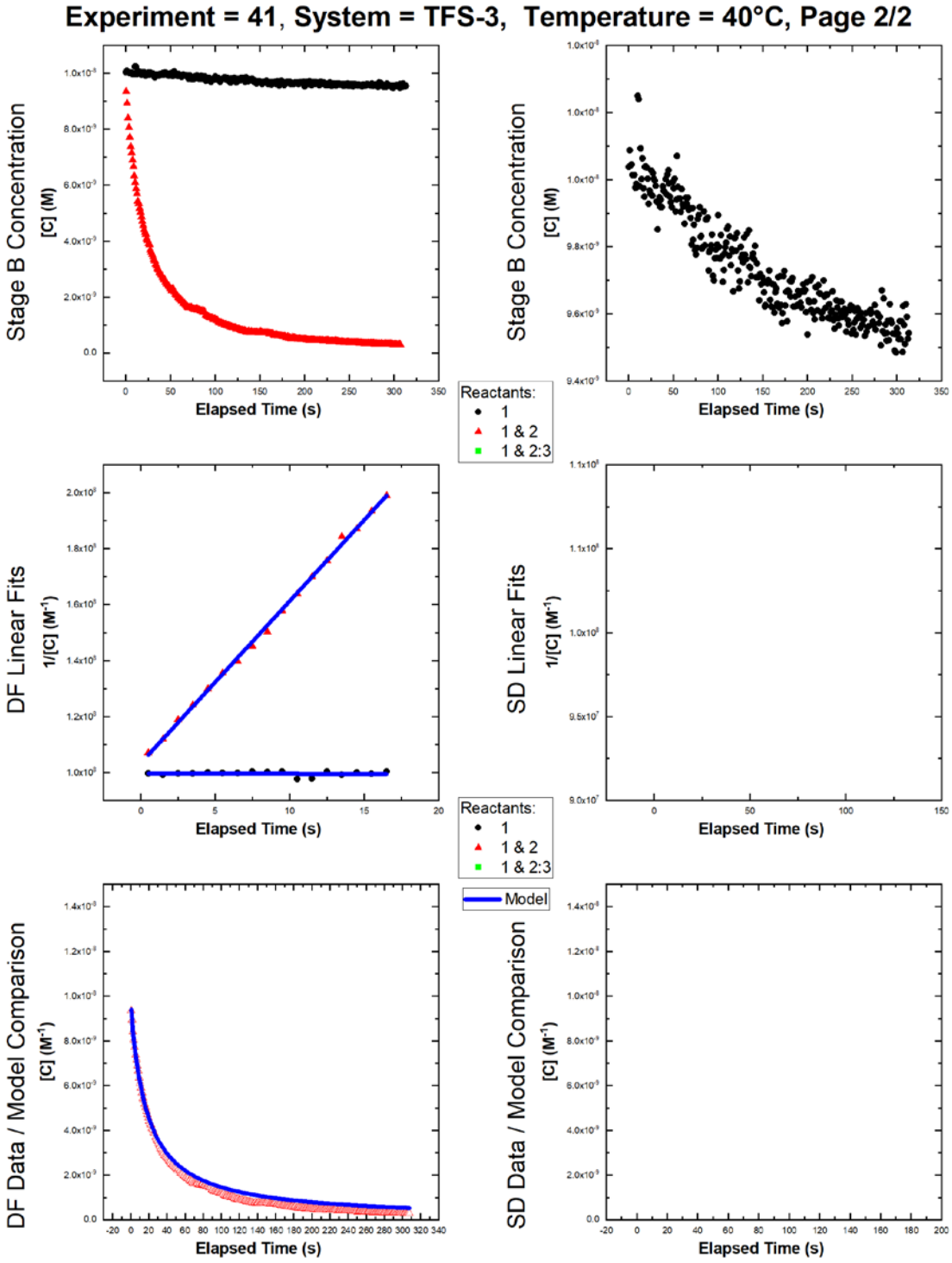

Experiment = 39, System = TFS-3, Temperature = 40°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

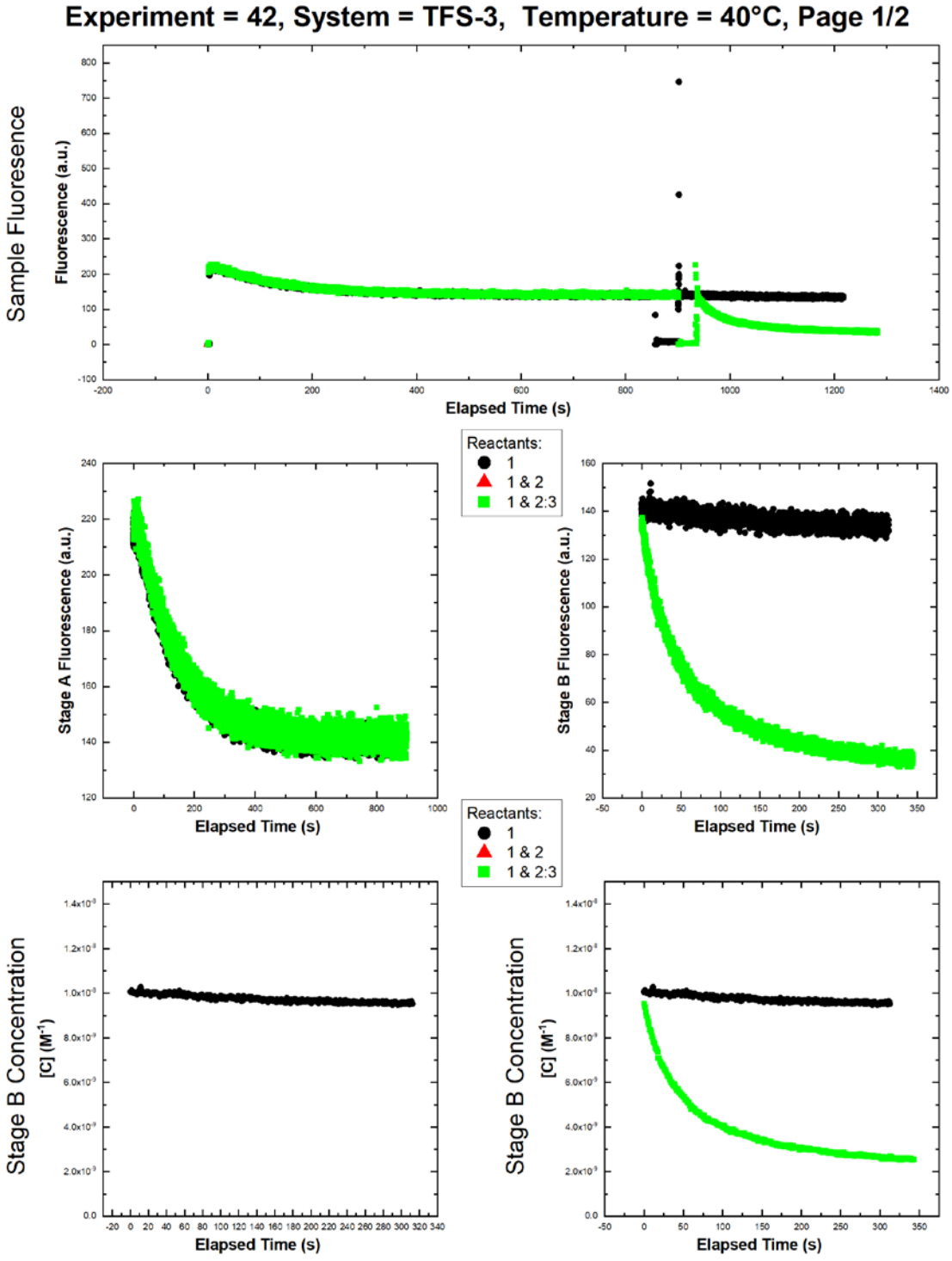

Experiment = 39, System = TFS-3, Temperature = 40°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

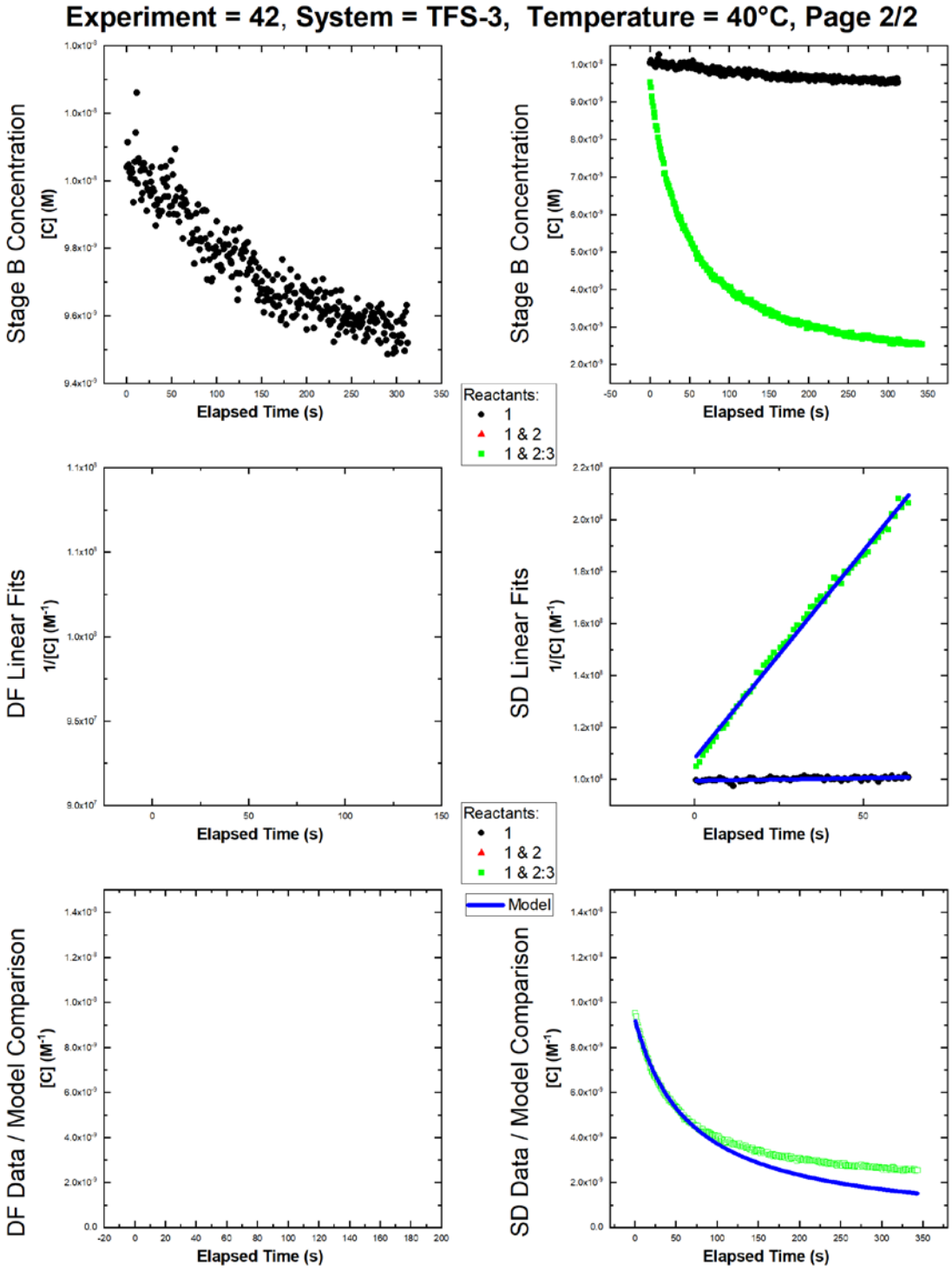

Experiment = 40, System = TFS-3, Temperature = 40°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

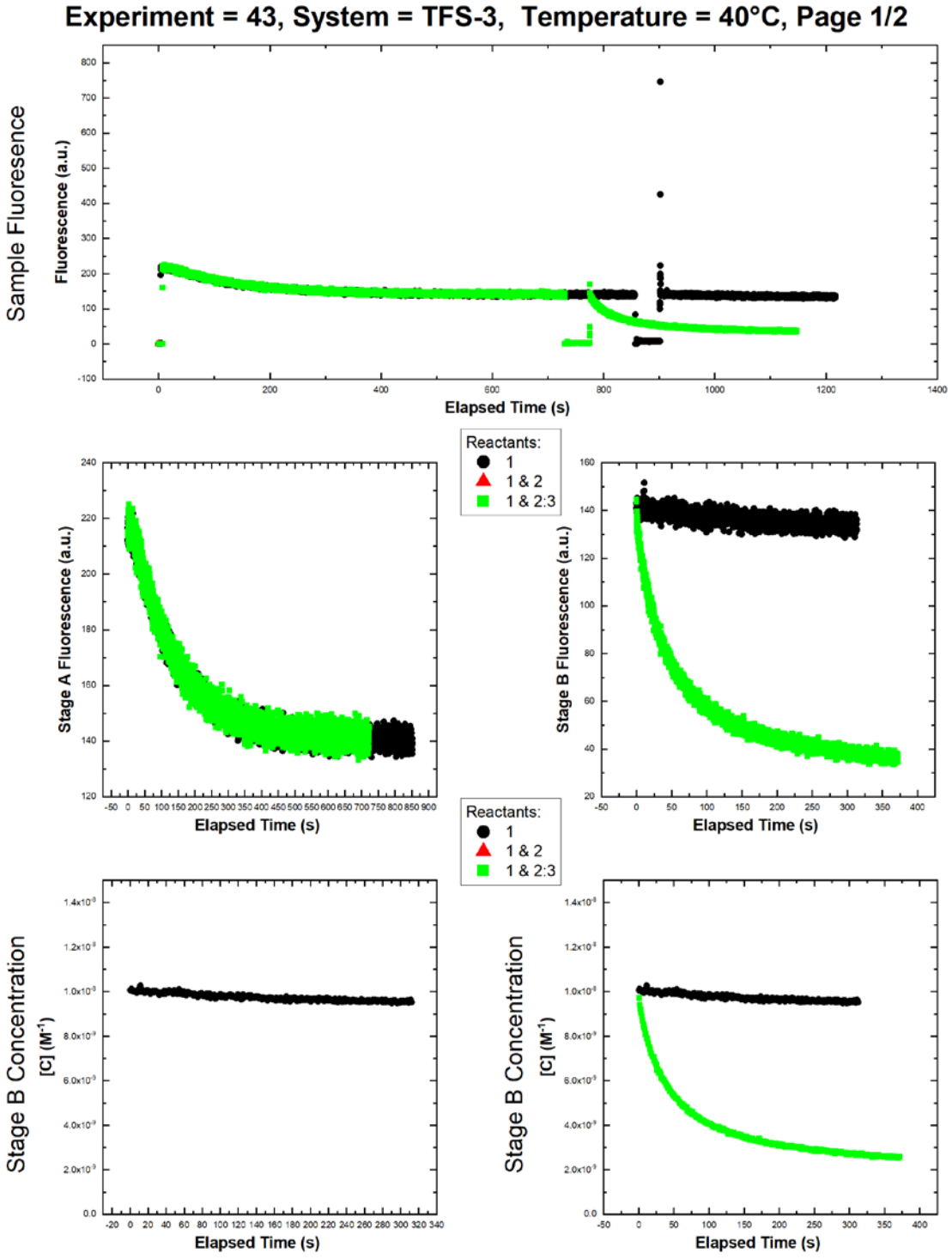

Experiment = 40, System = TFS-3, Temperature = 40°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
[C] (M-1)
SD Data / Model Comparison

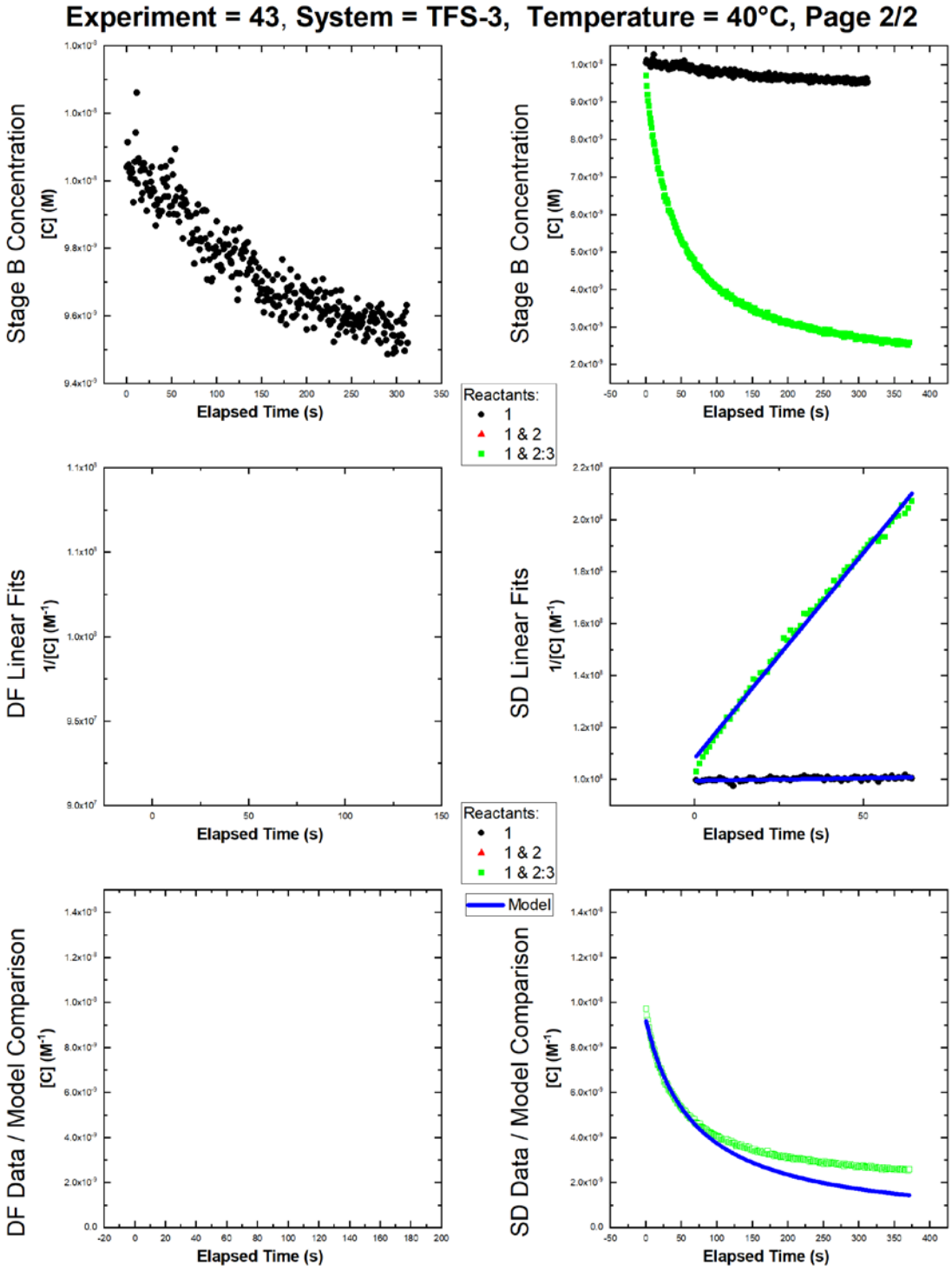

Experiment = 41, System = TFS-3, Temperature = 40°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

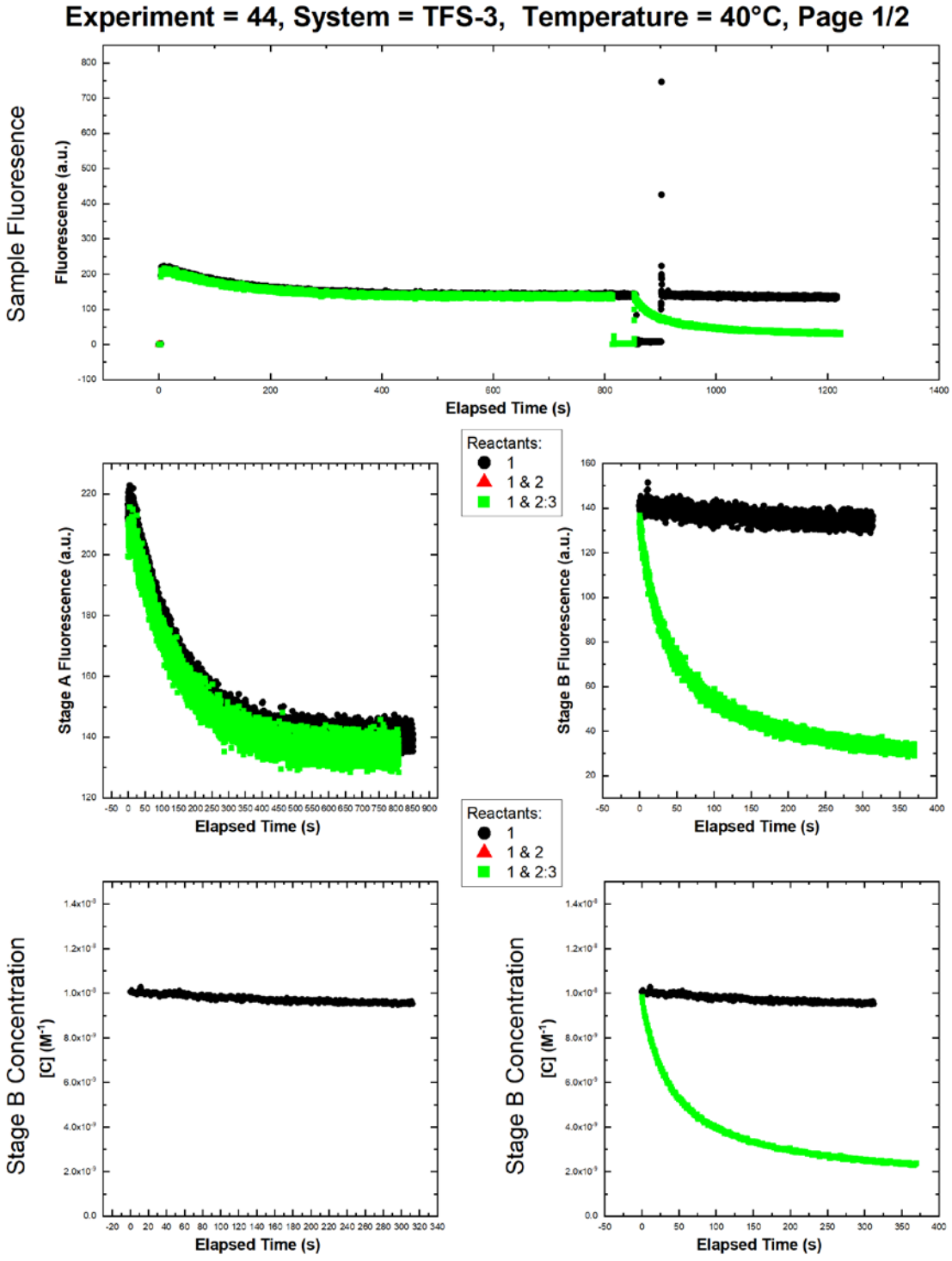

Experiment = 41, System = TFS-3, Temperature = 40°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

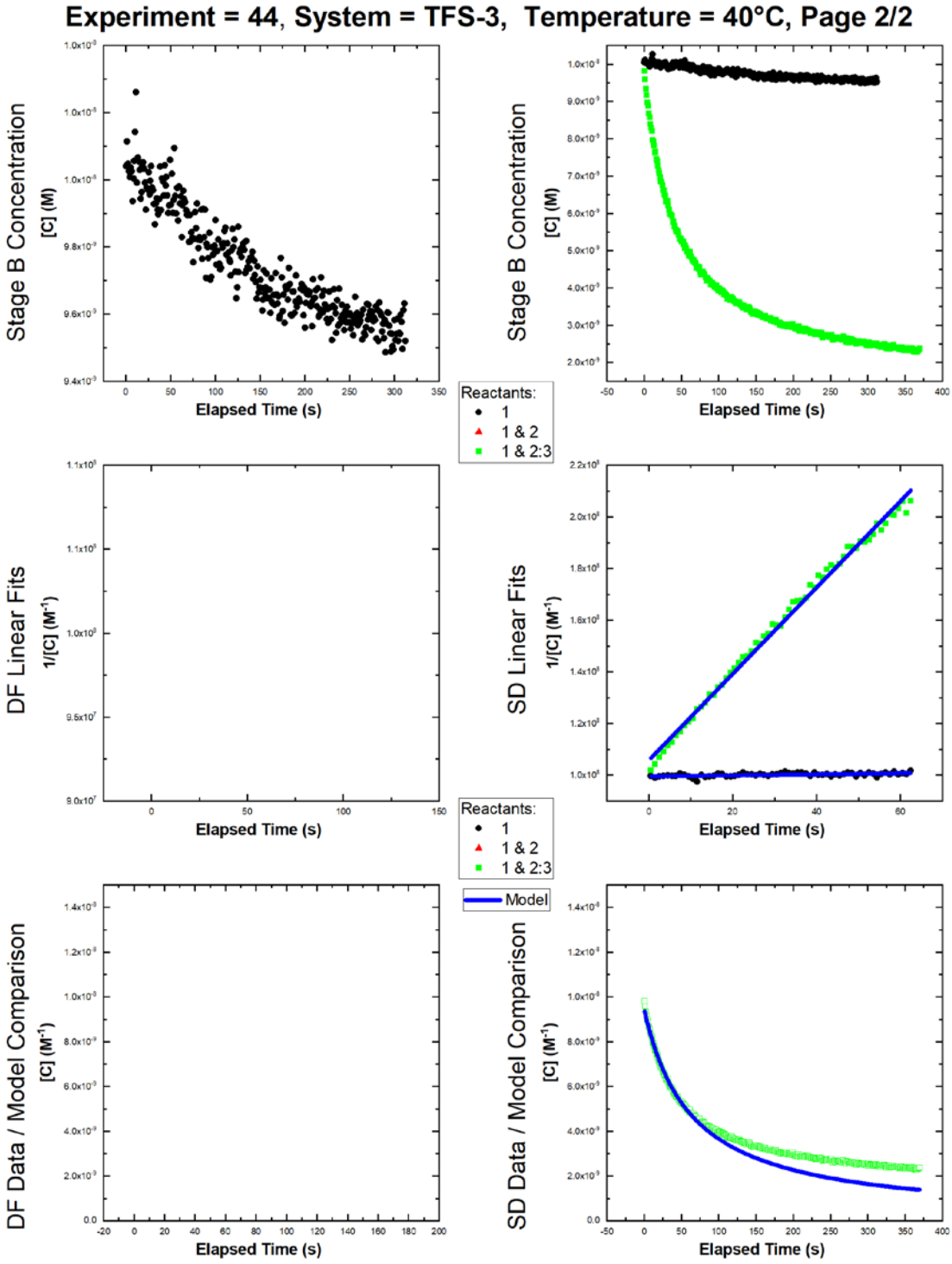

Experiment = 42, System = TFS-3, Temperature = 40°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

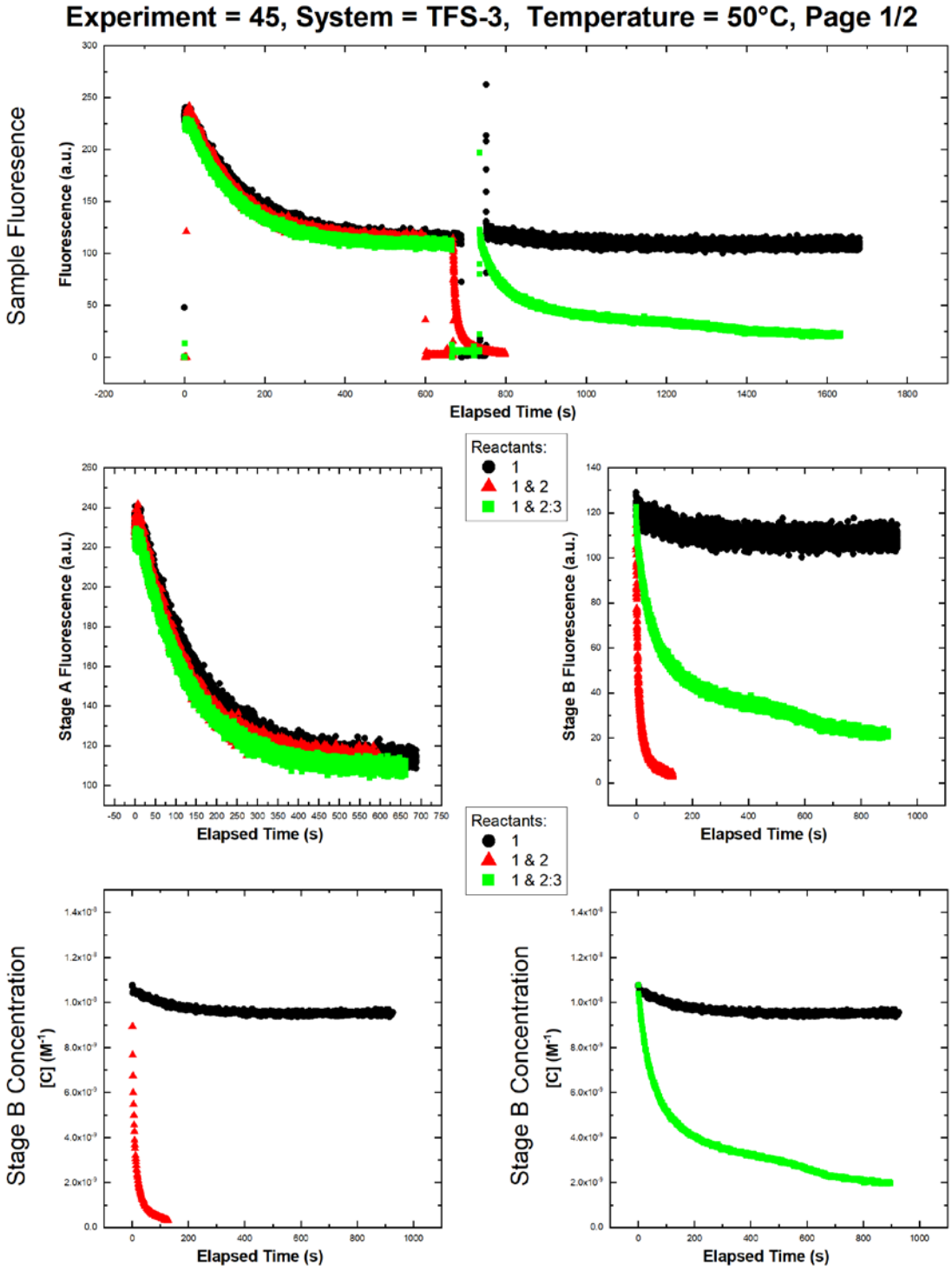

Experiment = 42, System = TFS-3, Temperature = 40°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

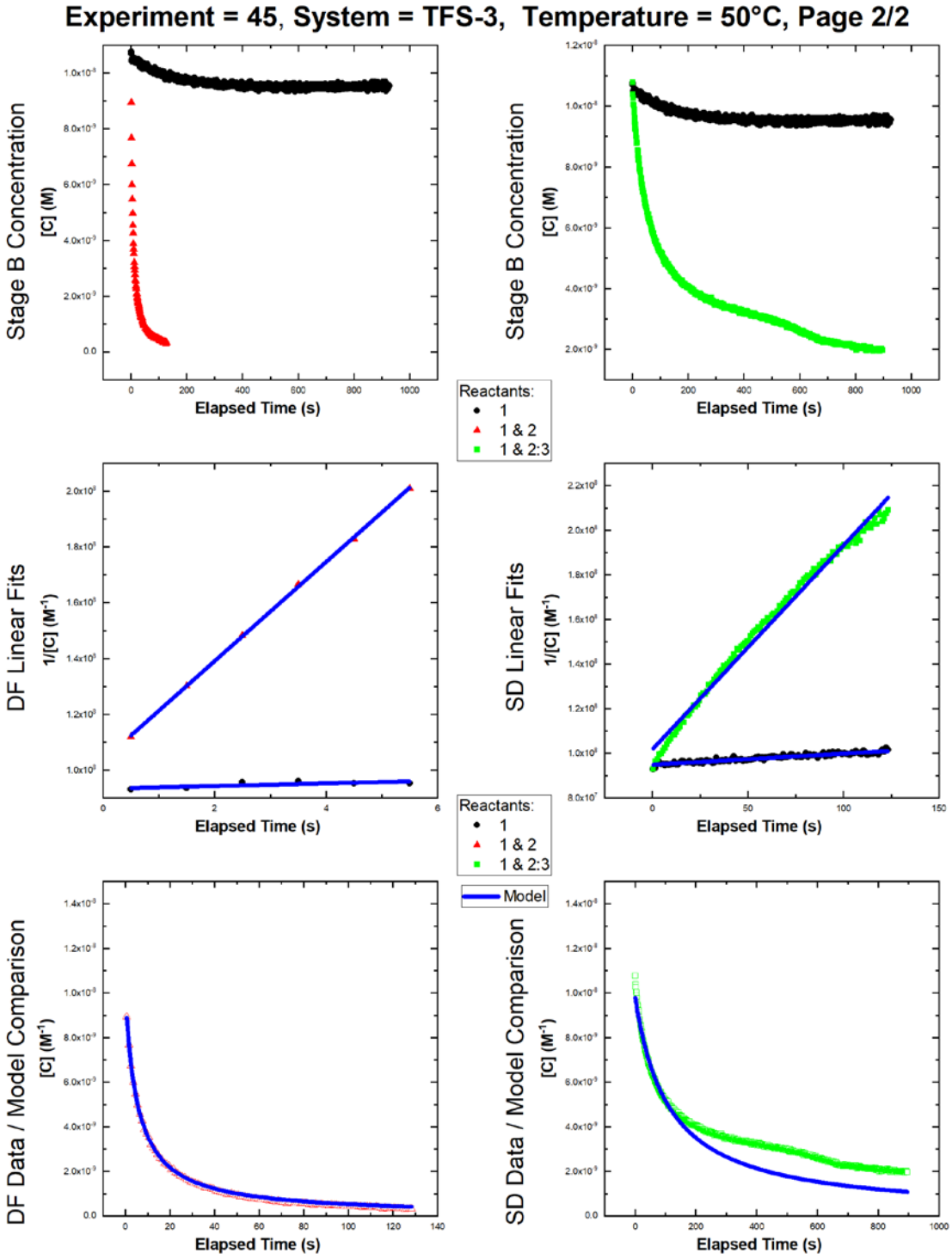

Experiment = 43, System = TFS-3, Temperature = 40°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

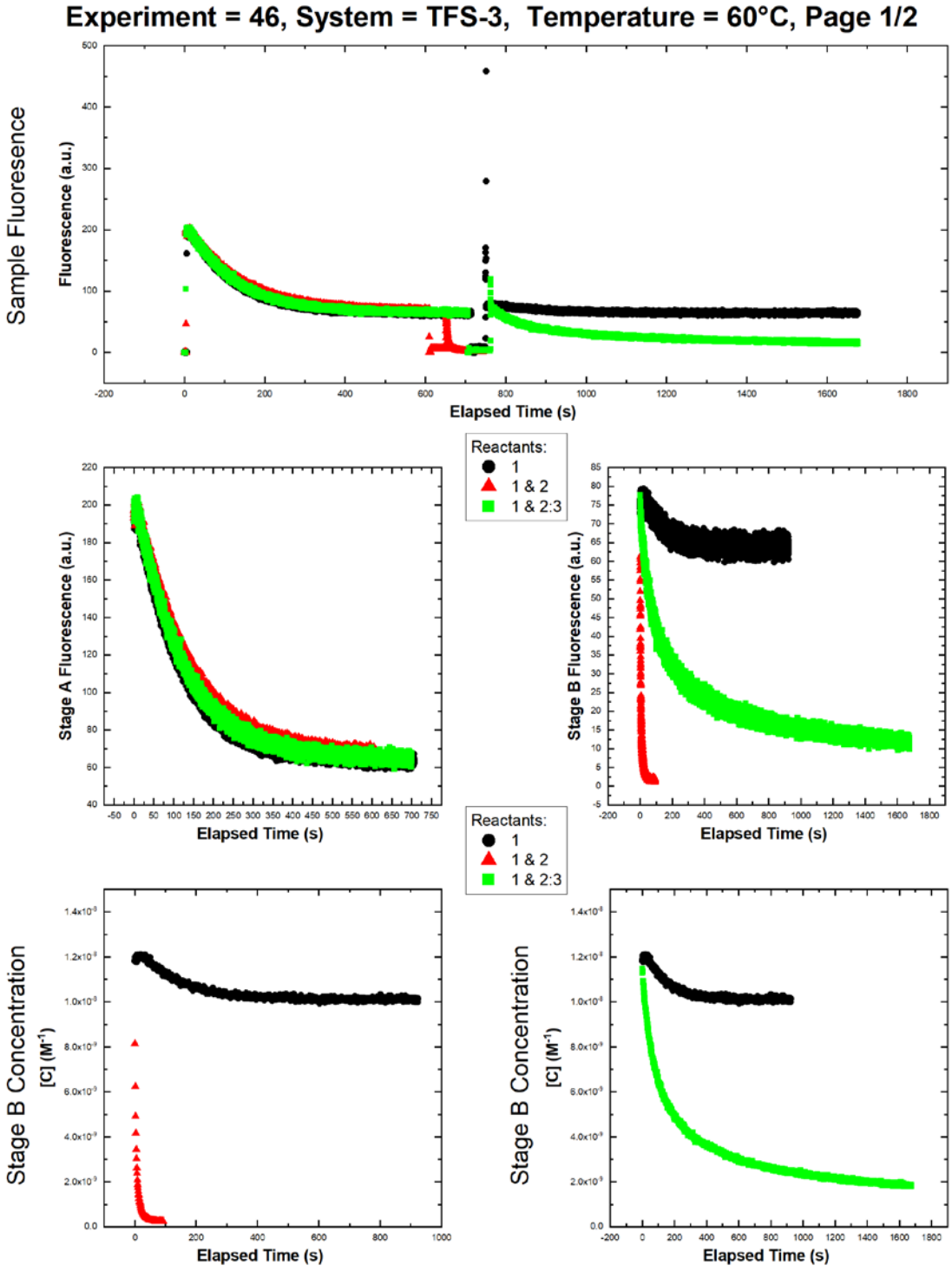

Experiment = 43, System = TFS-3, Temperature = 40°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

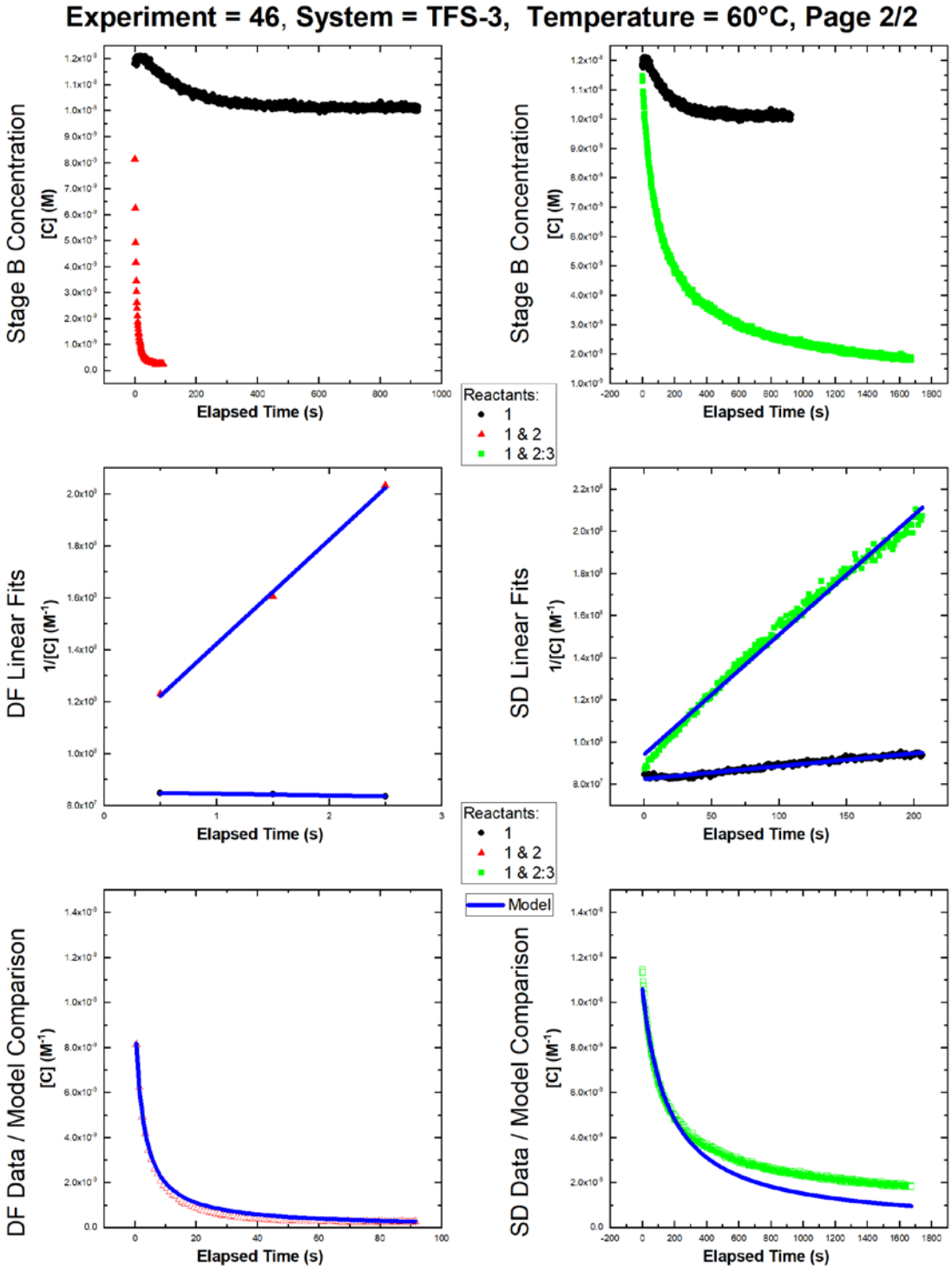

Experiment = 44, System = TFS-3, Temperature = 40°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

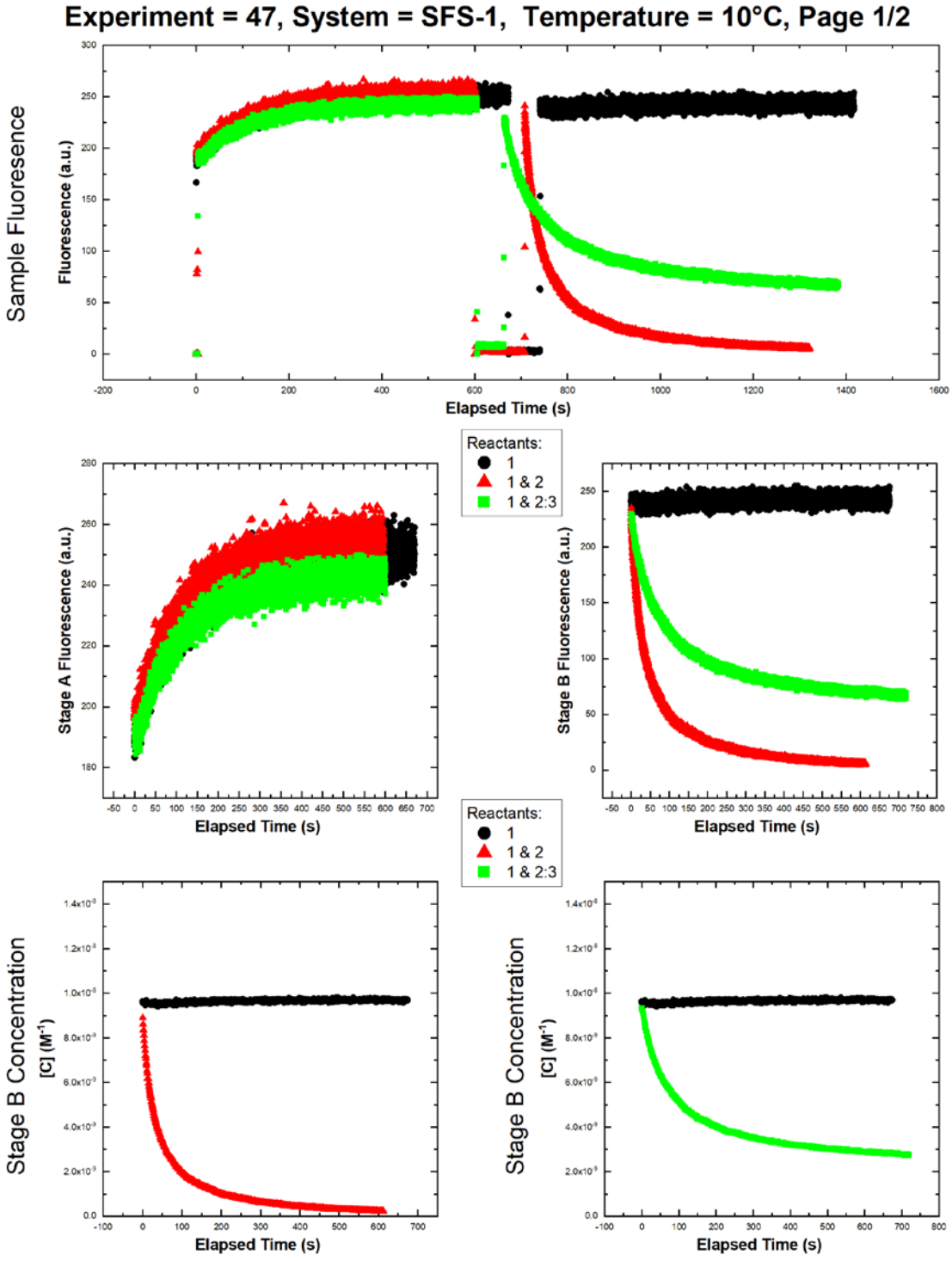

Experiment = 44, System = TFS-3, Temperature = 40°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

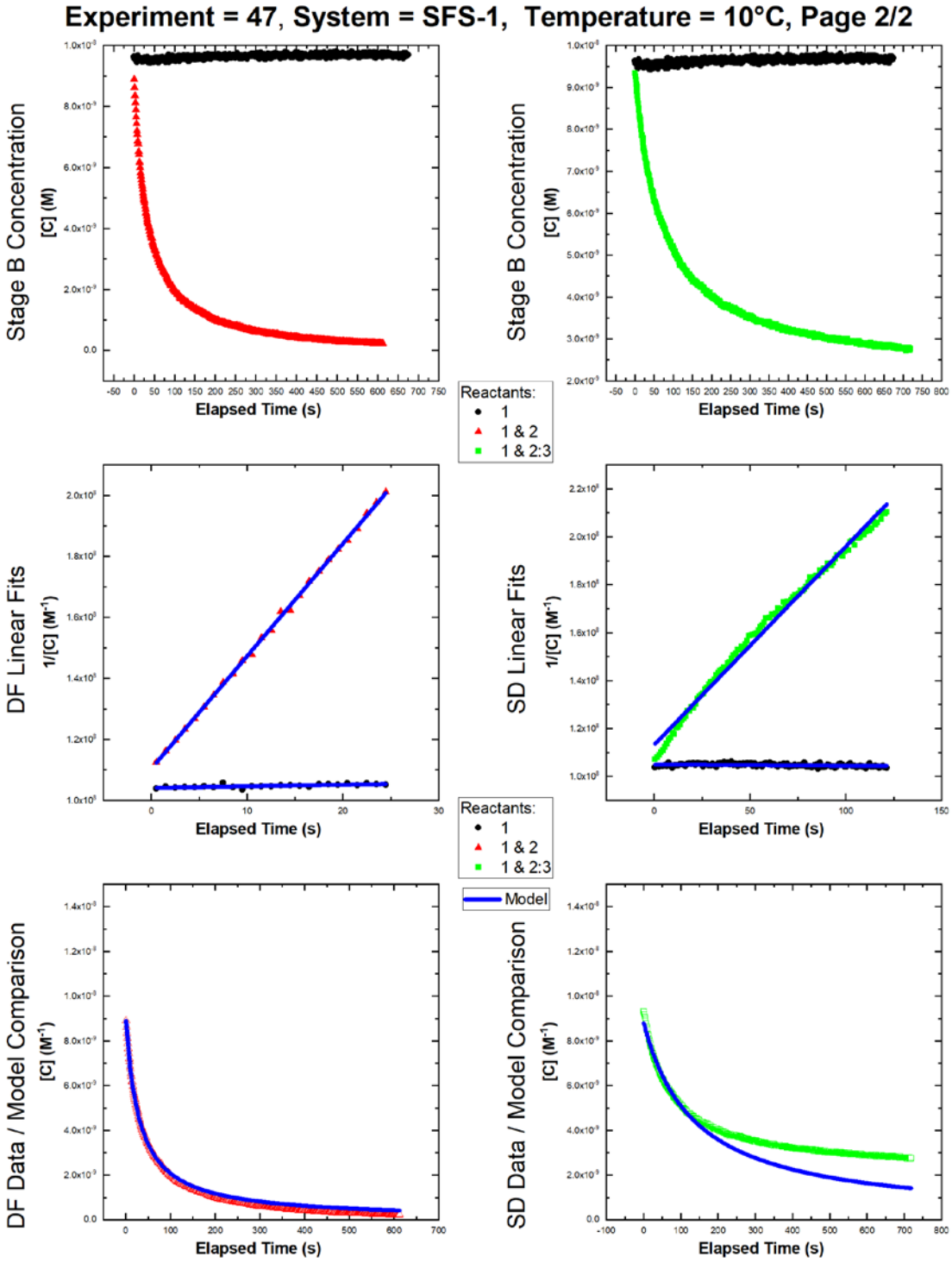

Experiment = 45, System = TFS-3, Temperature = 50°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

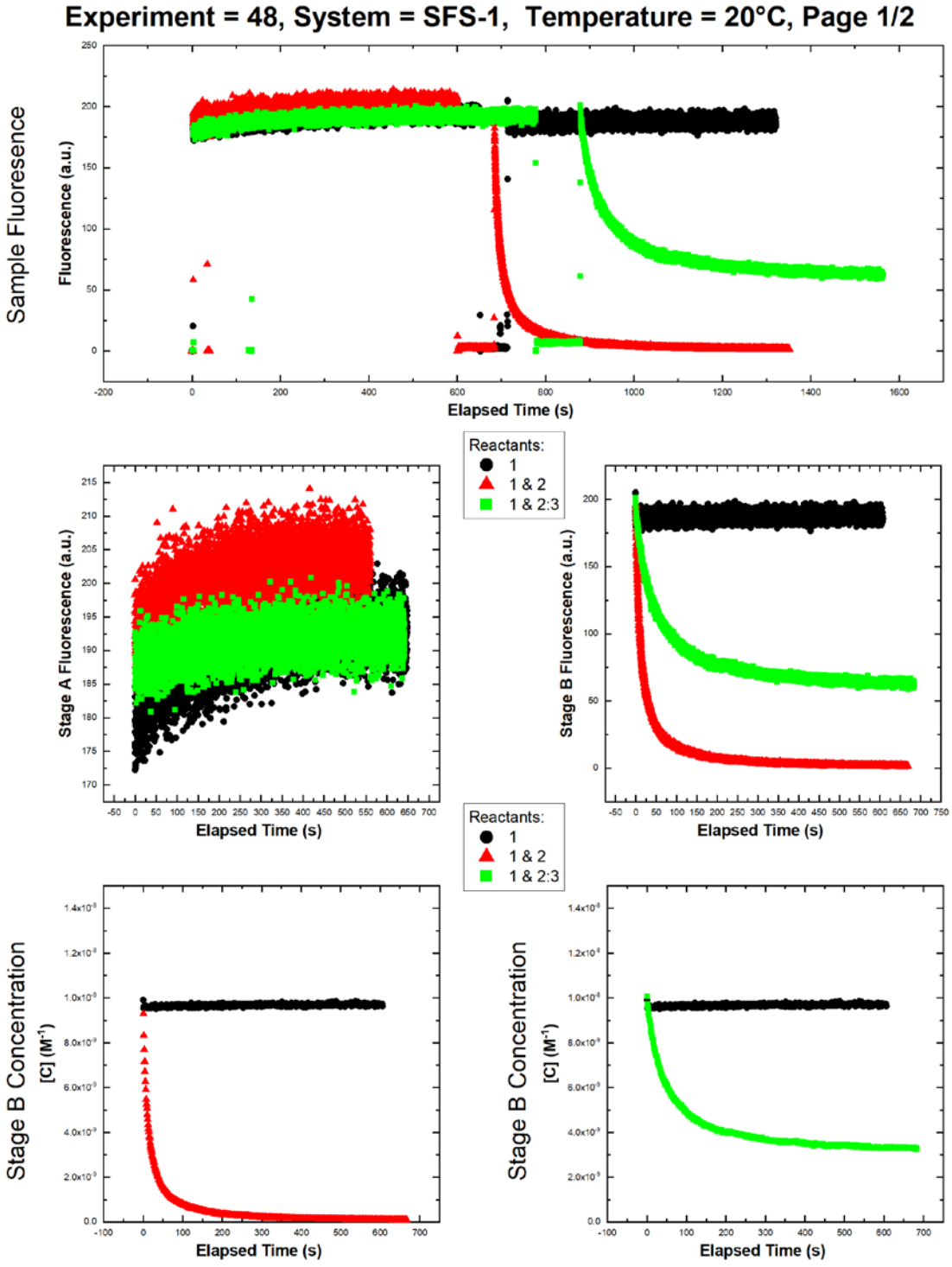

Experiment = 45, System = TFS-3, Temperature = 50°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

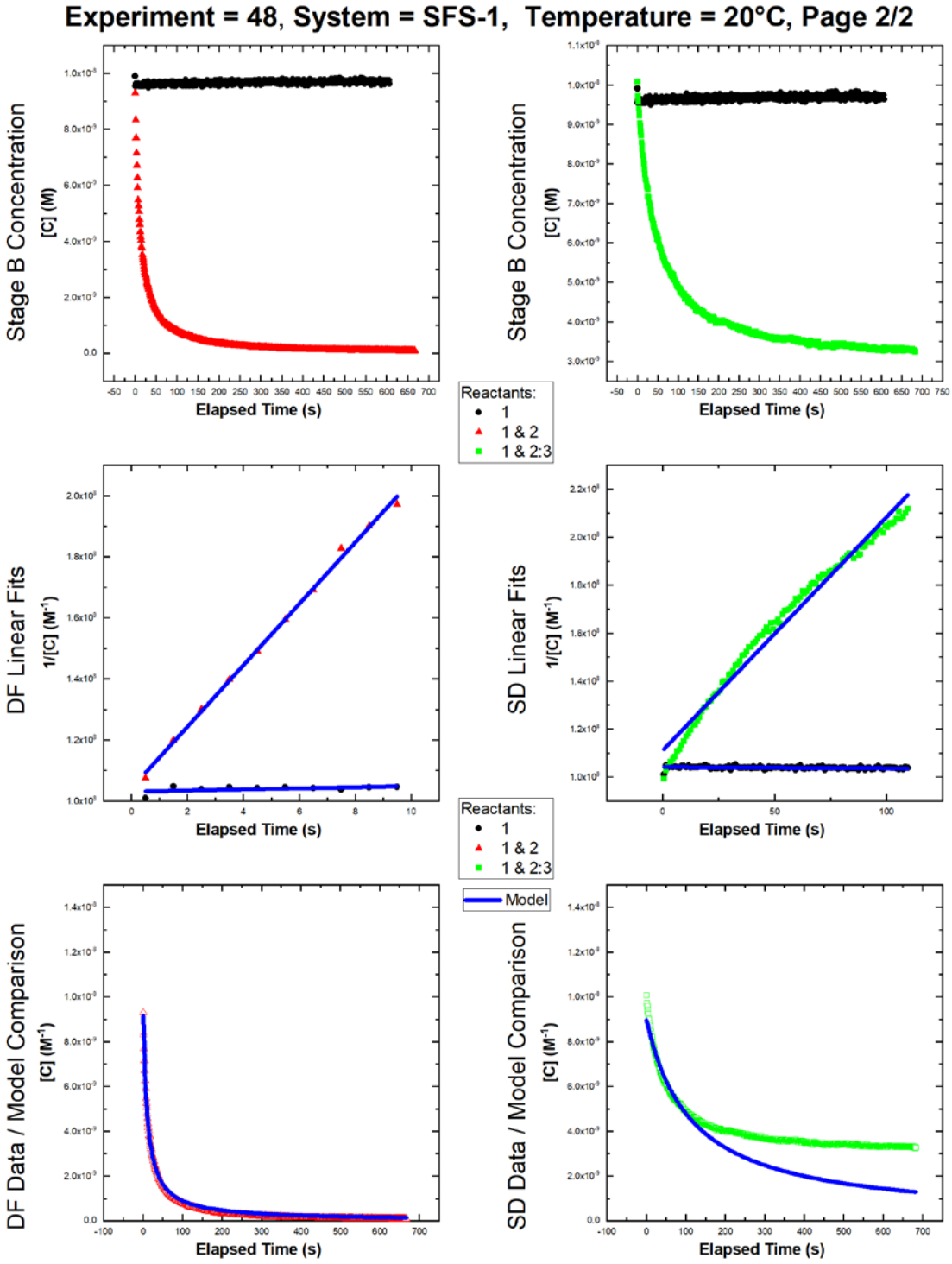

Experiment = 46, System = TFS-3, Temperature = 60°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

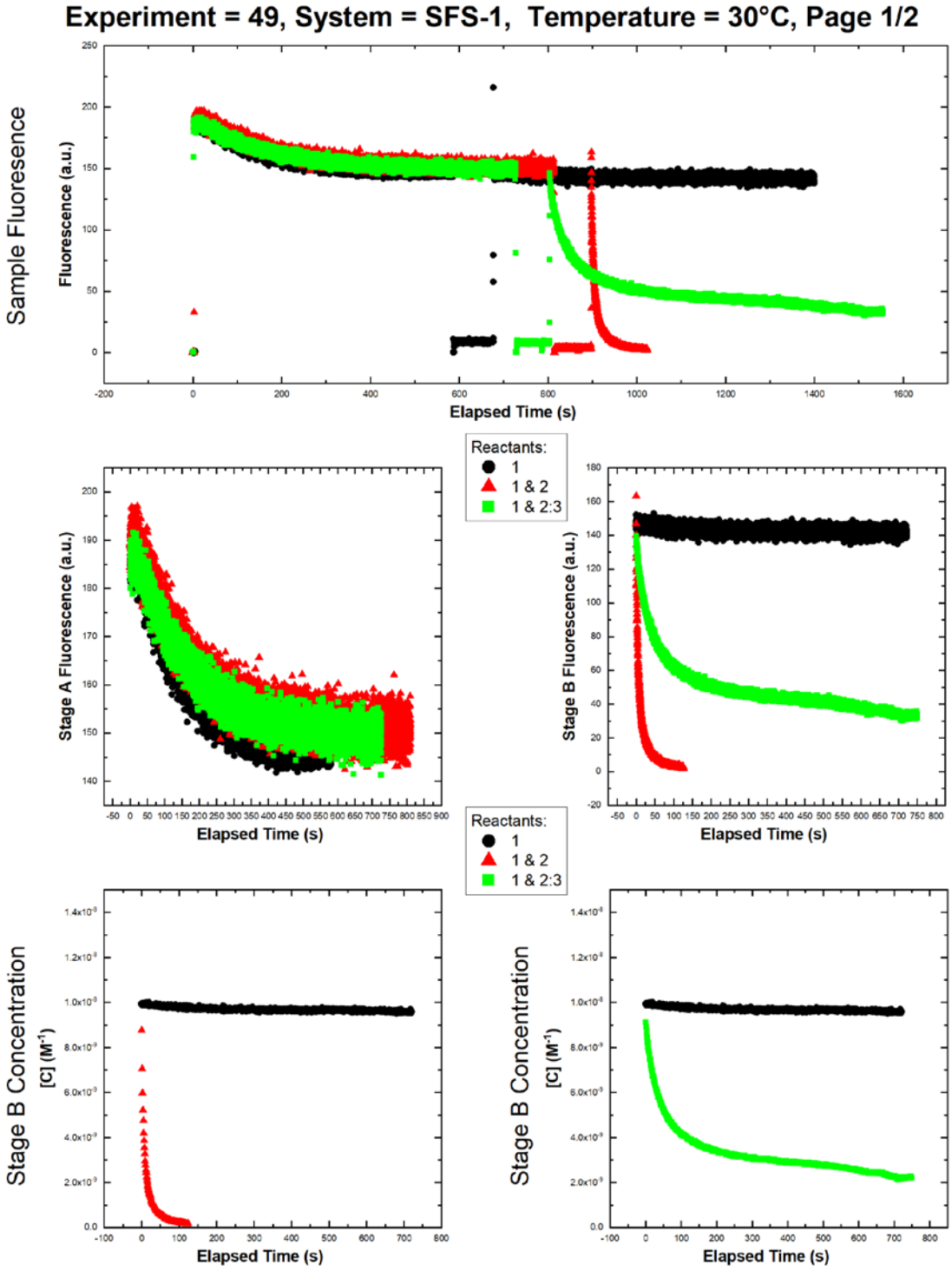

Experiment = 46, System = TFS-3, Temperature = 60°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

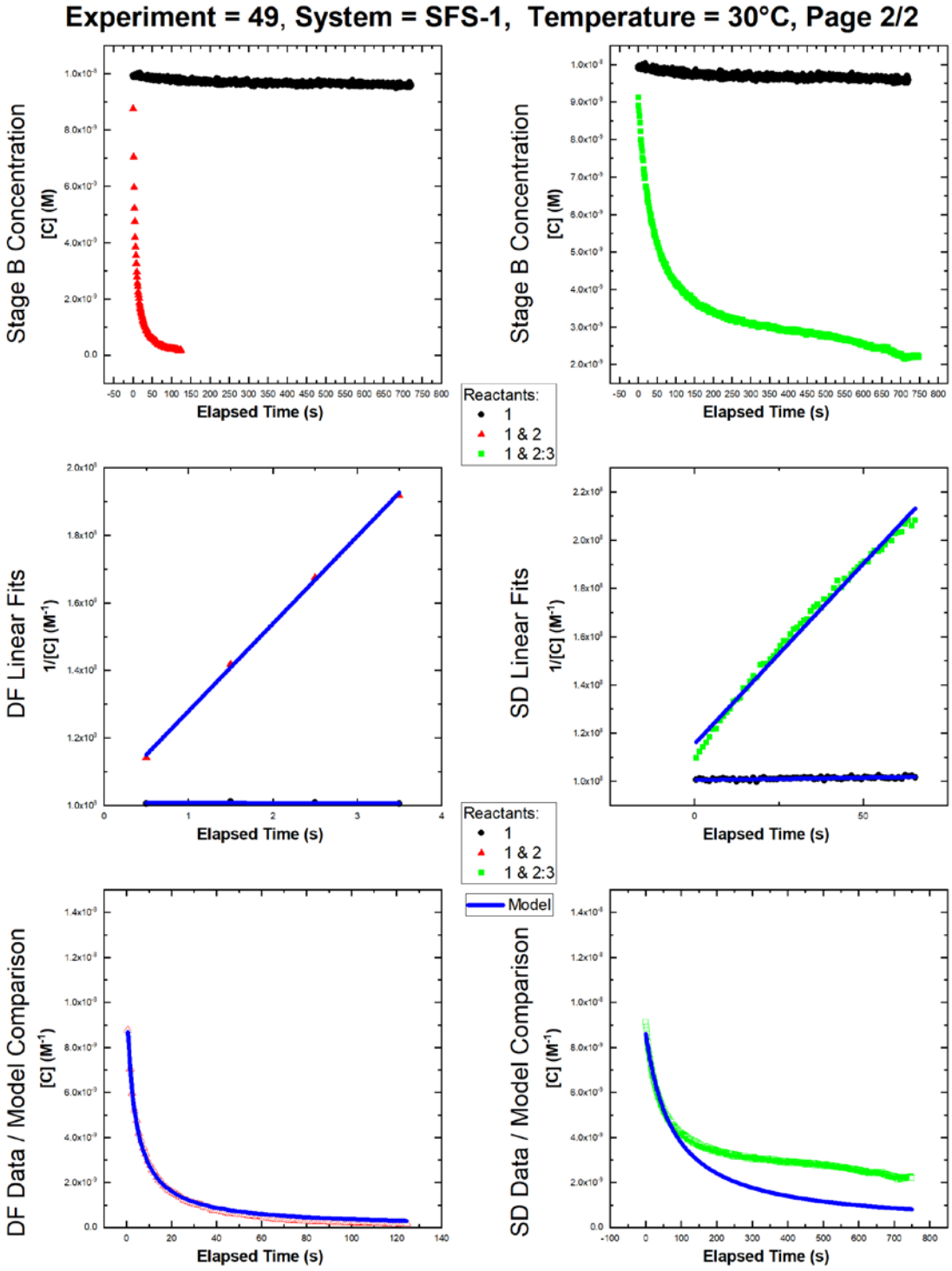

Experiment = 47, System = SFS-1, Temperature = 10°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

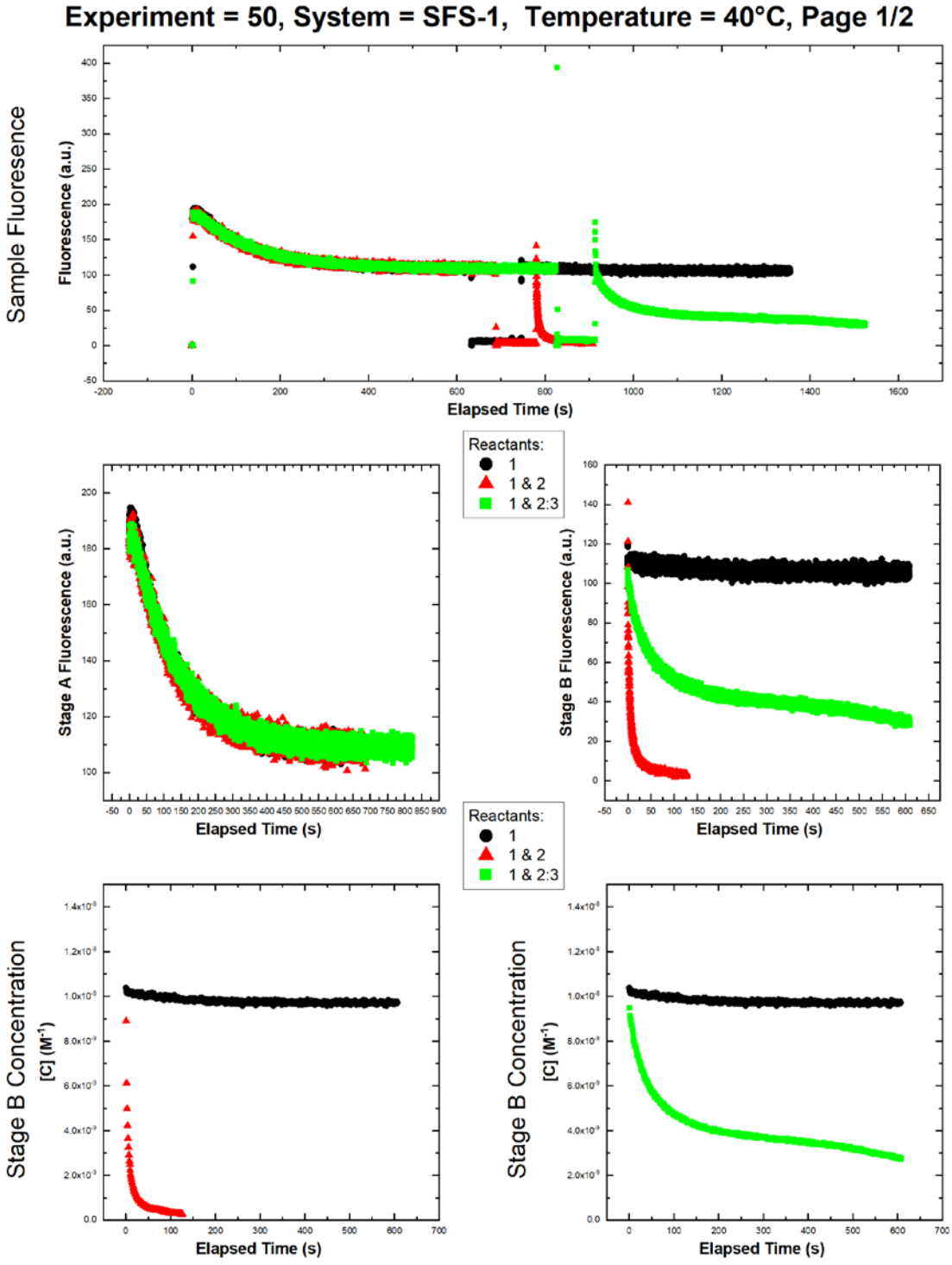

Experiment = 47, System = SFS-1, Temperature = 10°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

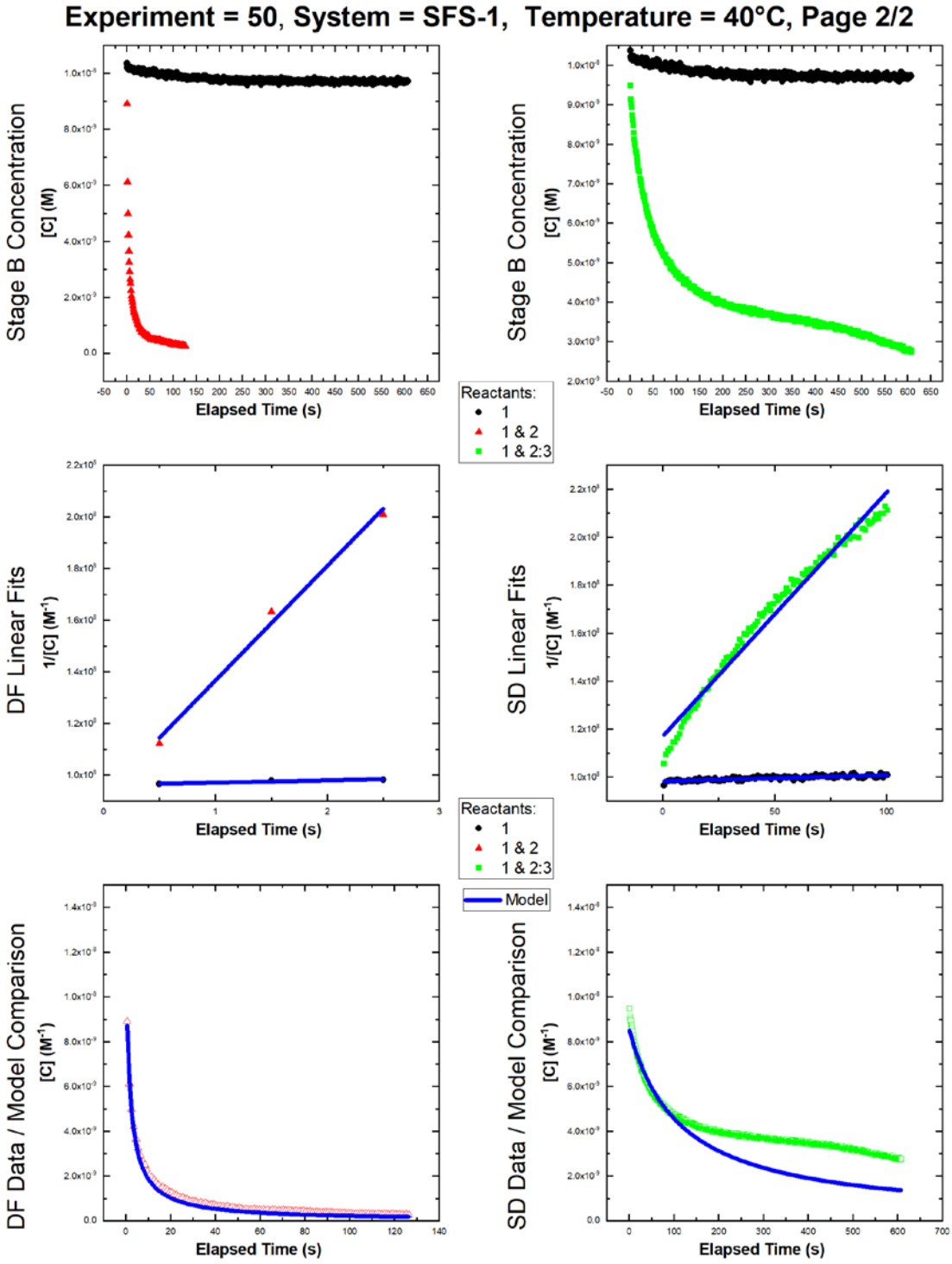

Experiment = 48, System = SFS-1, Temperature = 20°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

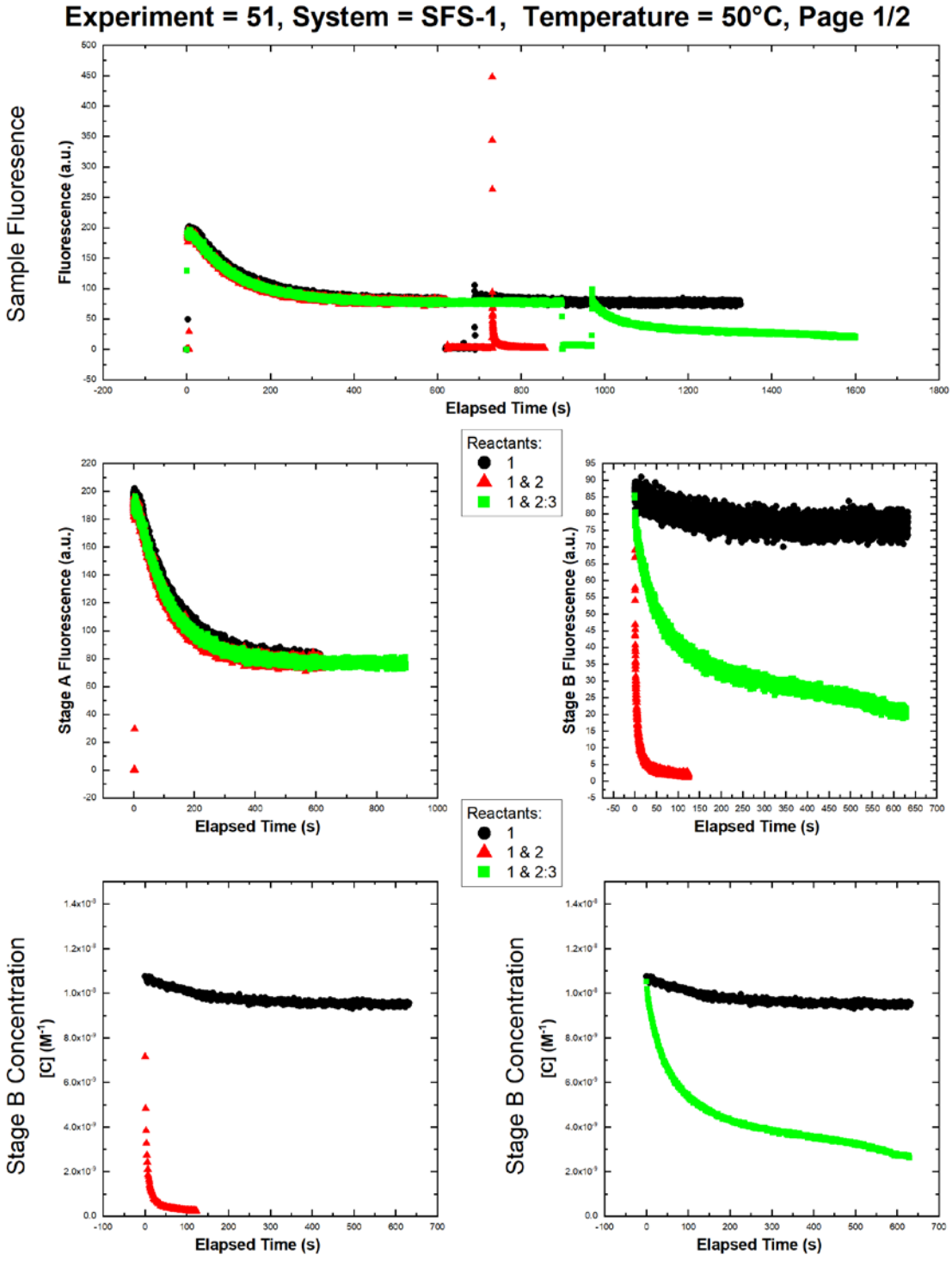

Experiment = 48, System = SFS-1, Temperature = 20°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

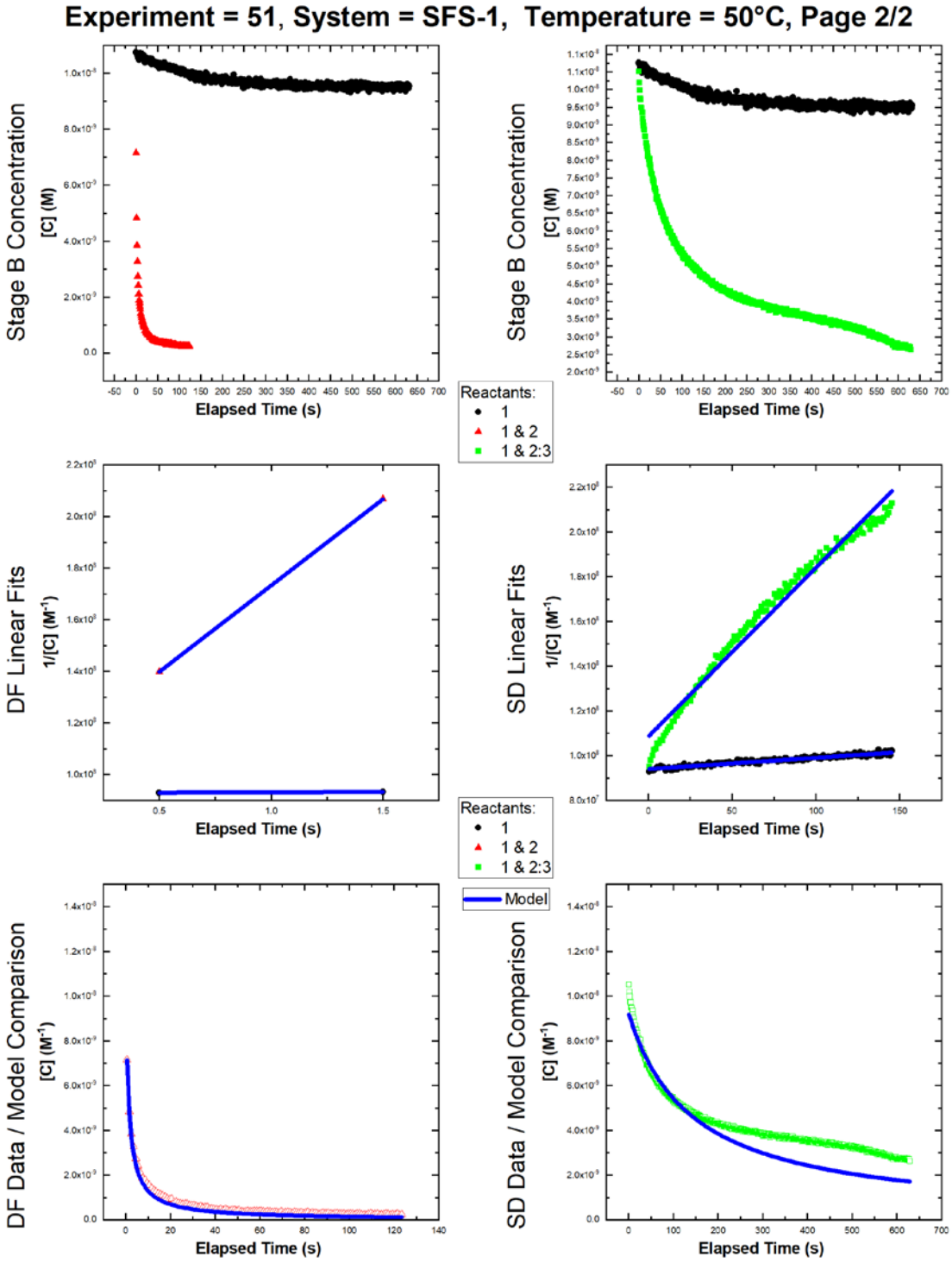

Experiment = 49, System = SFS-1, Temperature = 30°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

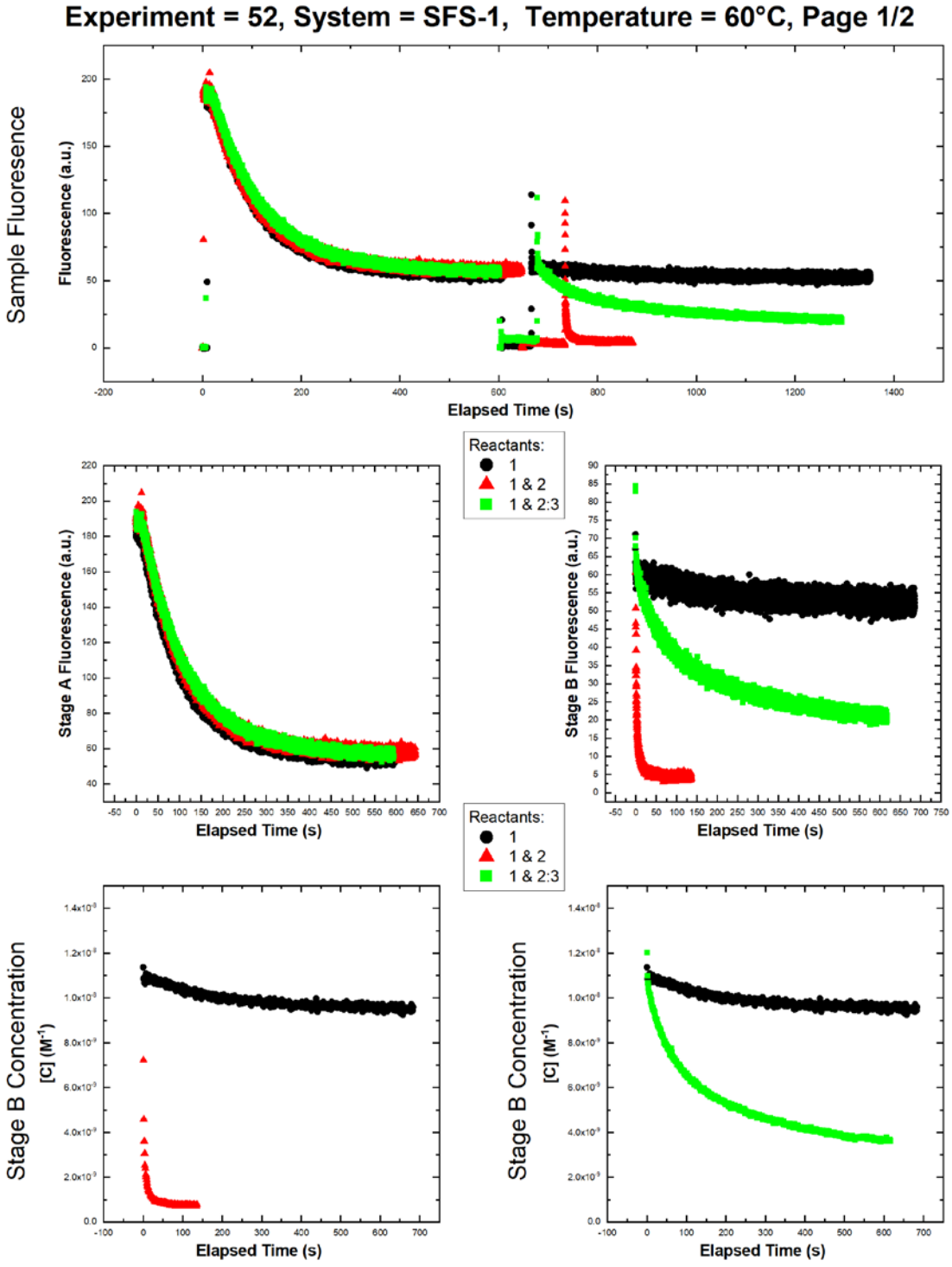

Experiment = 49, System = SFS-1, Temperature = 30°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

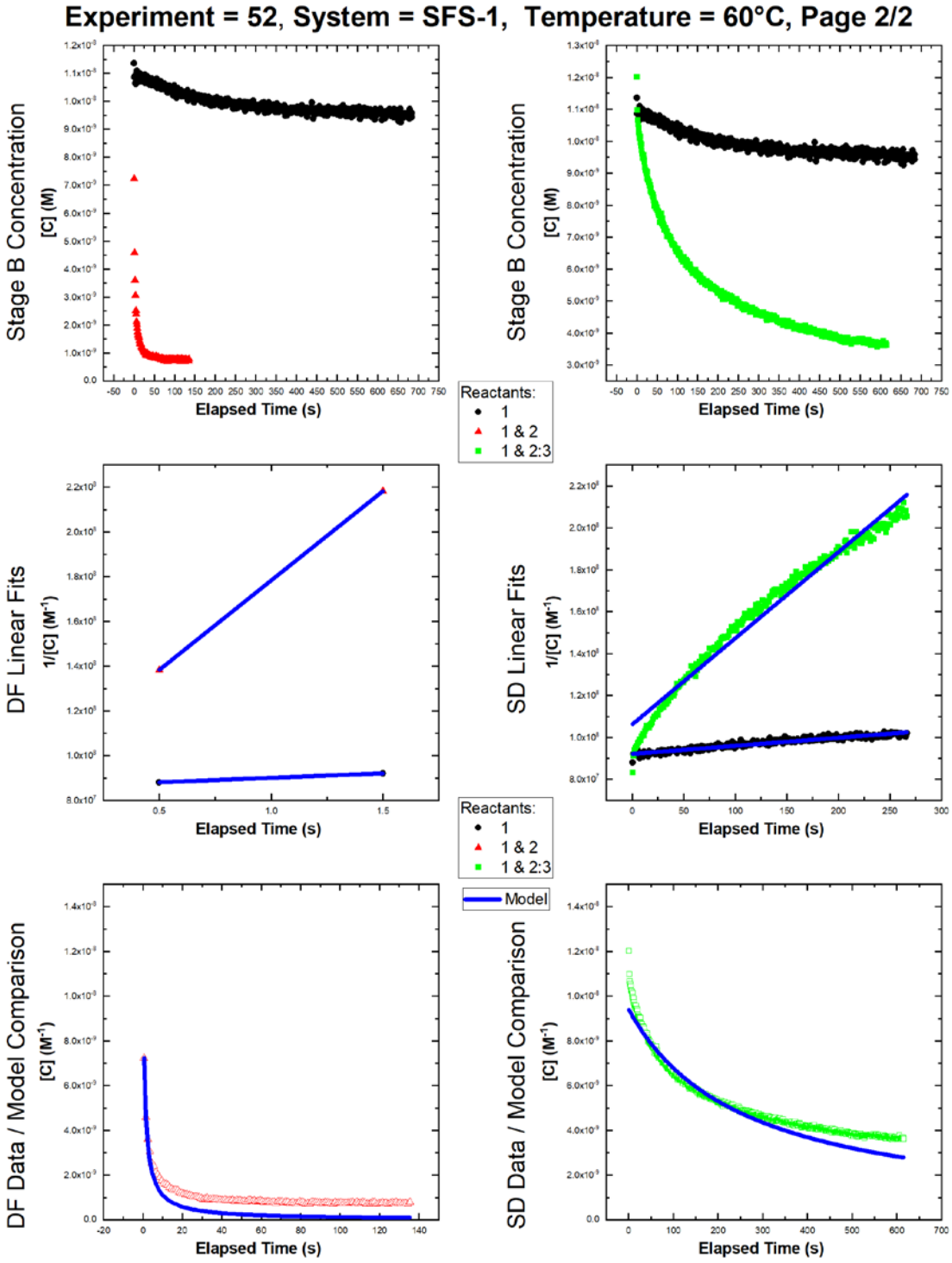

Experiment = 50, System = SFS-1, Temperature = 40°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

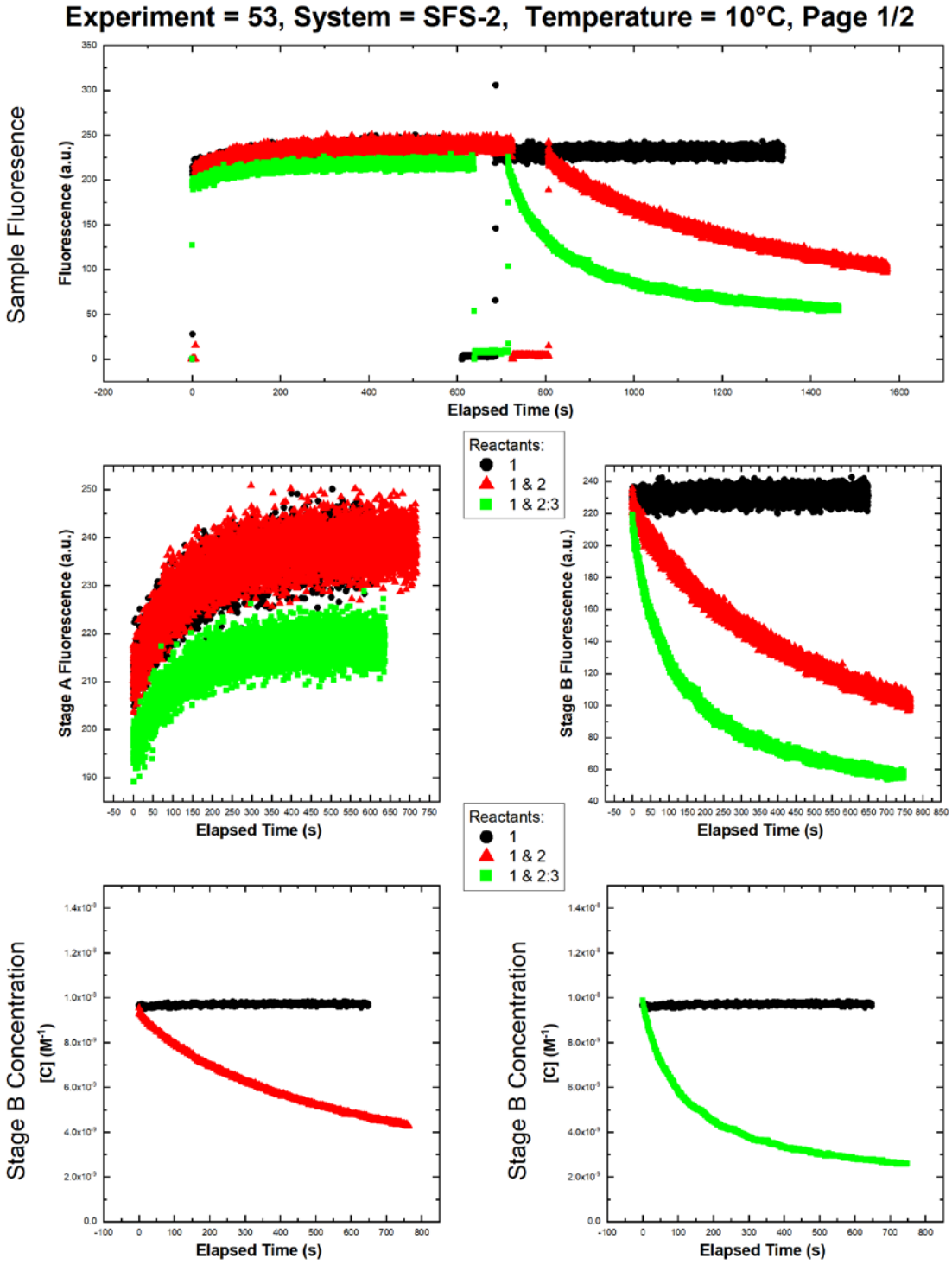

Experiment = 50, System = SFS-1, Temperature = 40°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

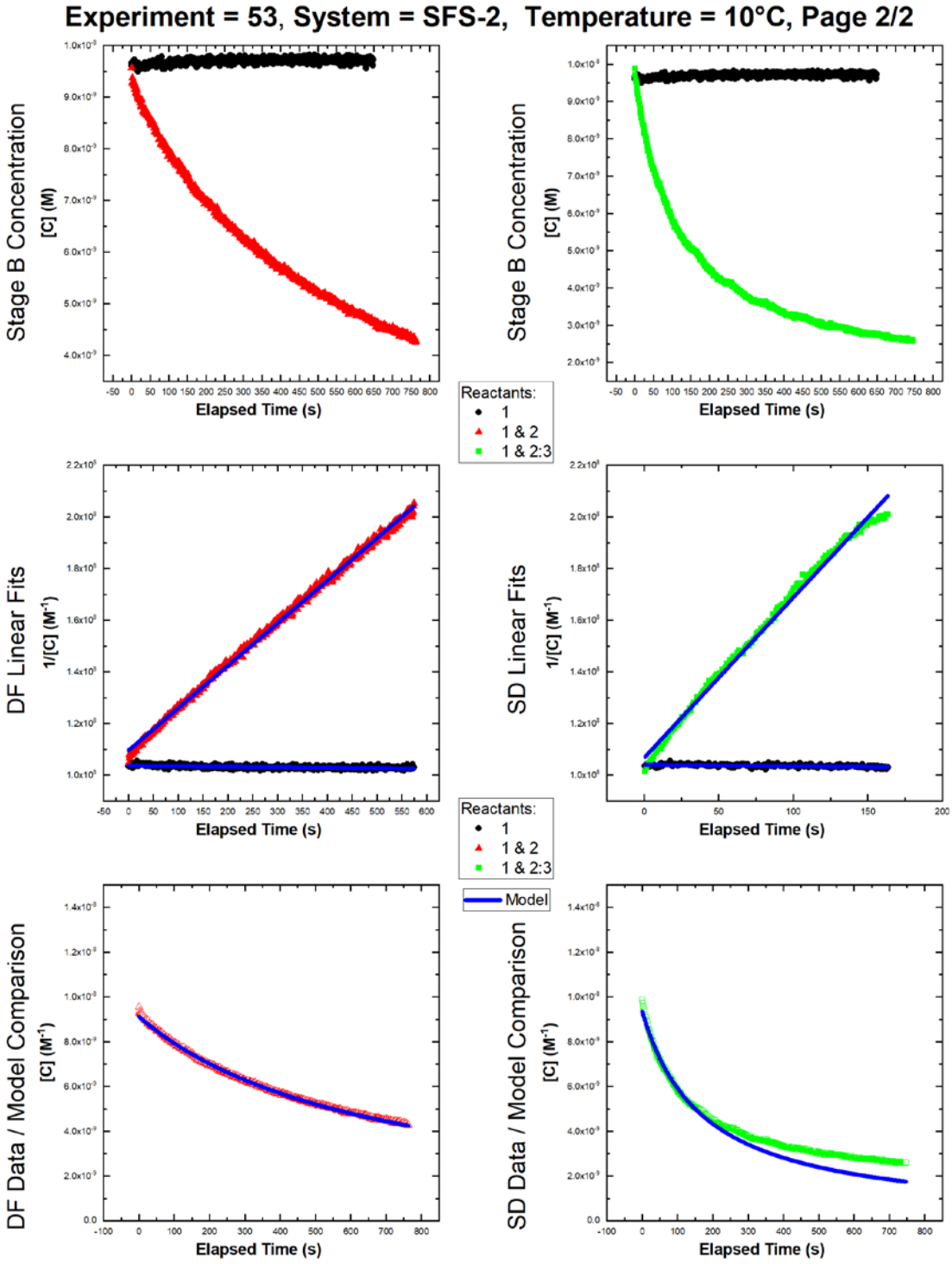

Experiment = 51, System = SFS-1, Temperature = 50°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

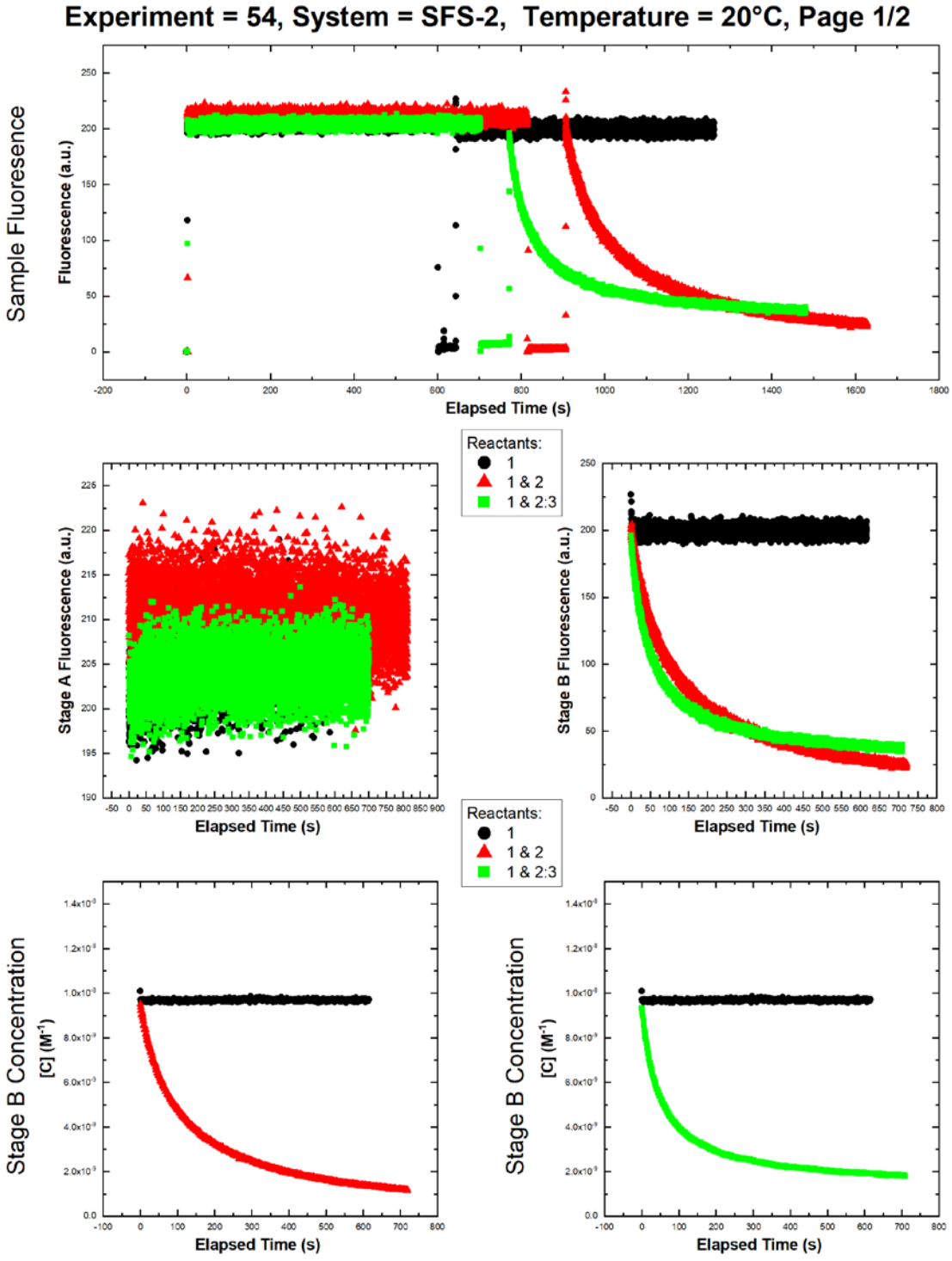

Experiment = 51, System = SFS-1, Temperature = 50°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

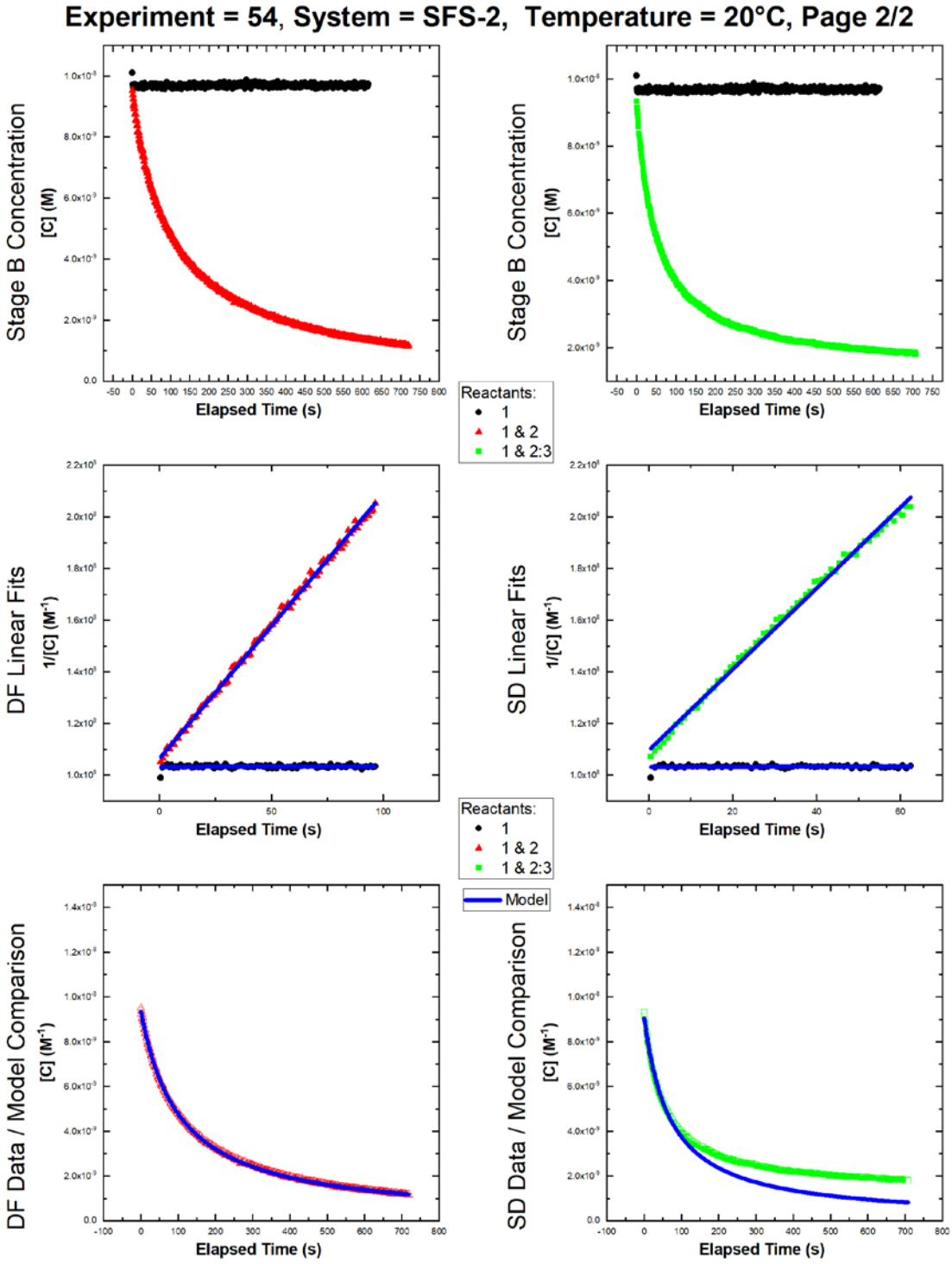

Experiment = 52, System = SFS-1, Temperature = 60°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

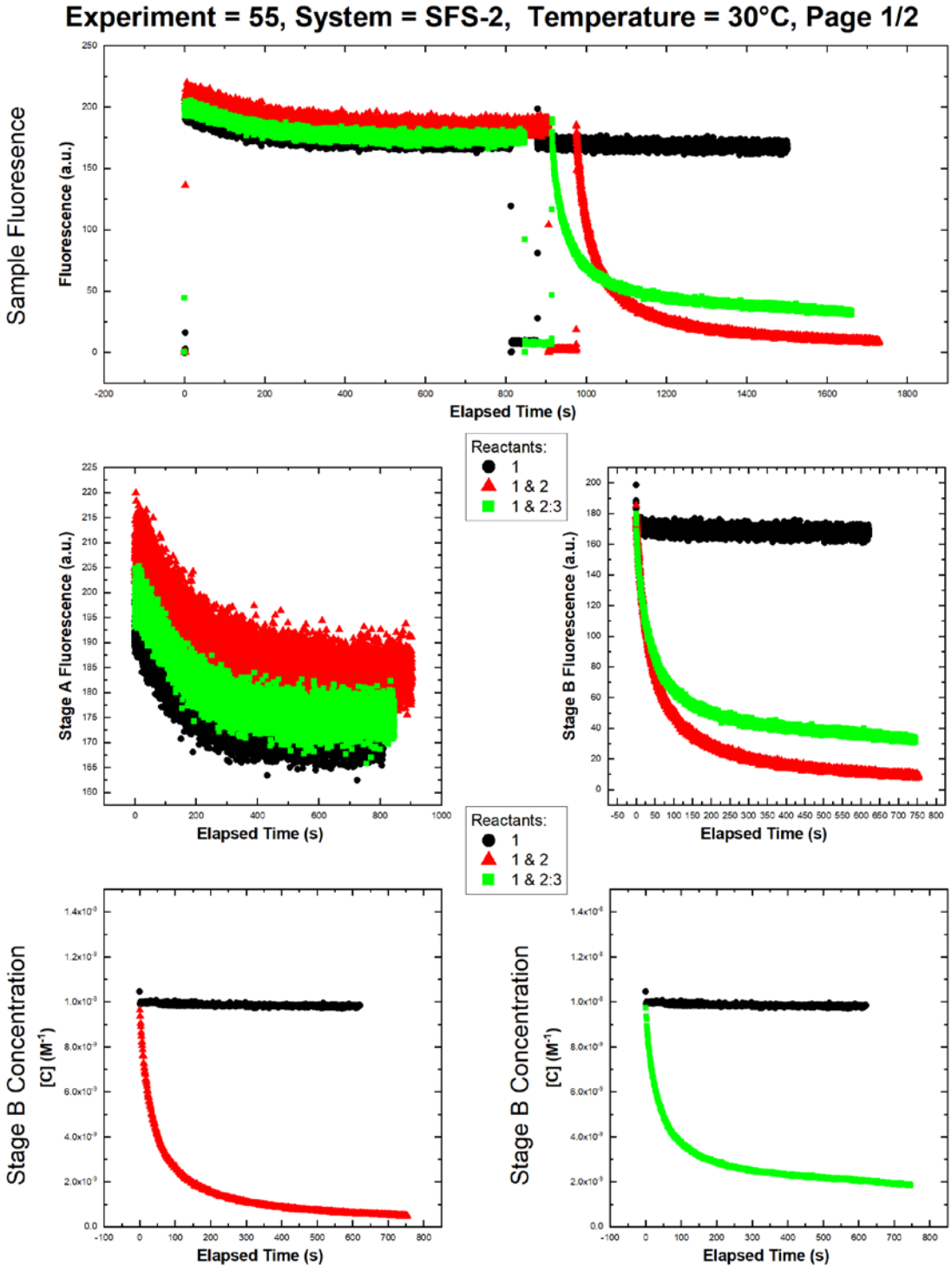

Experiment = 52, System = SFS-1, Temperature = 60°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

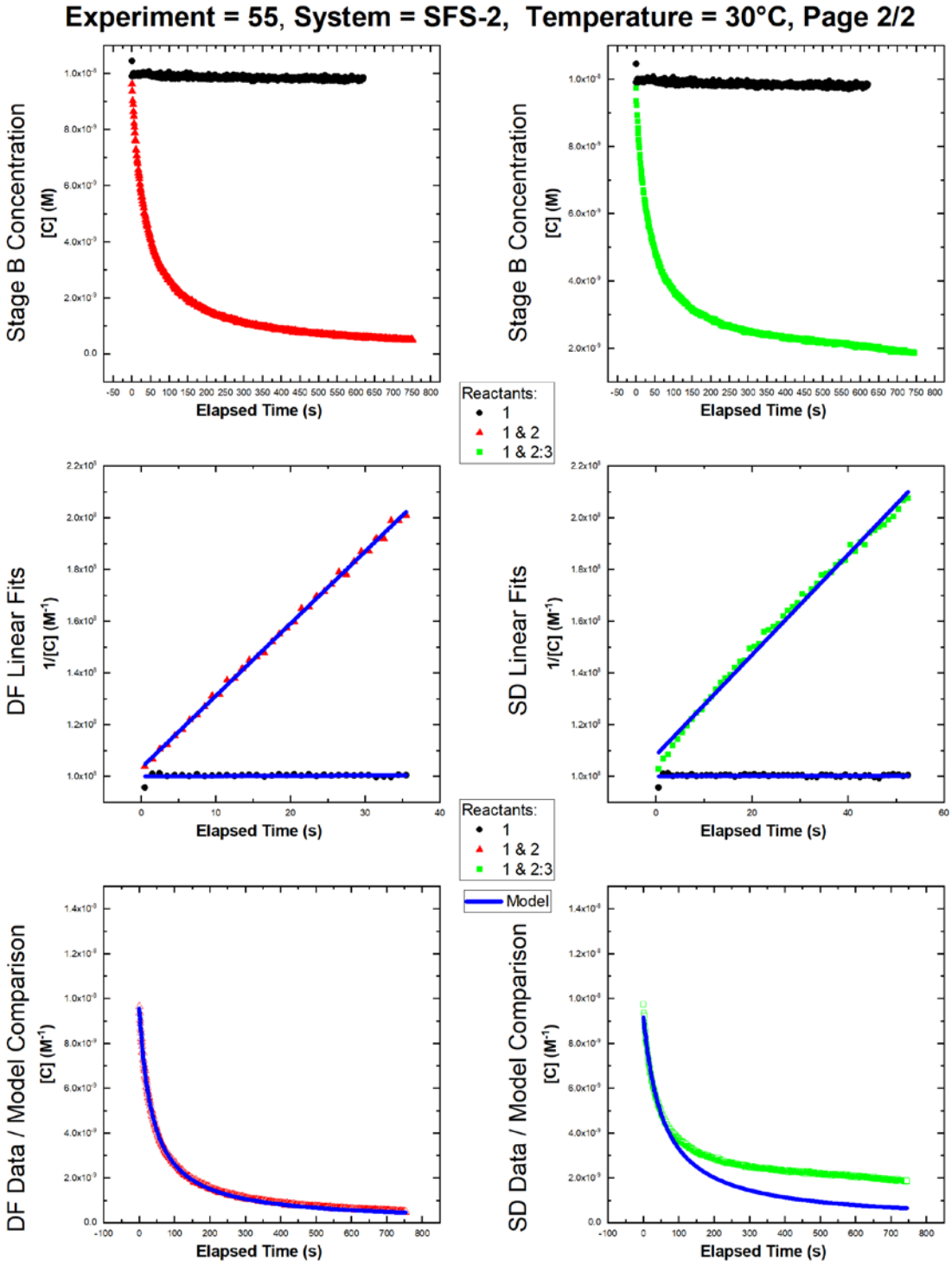

Experiment = 53, System = SFS-2, Temperature = 10°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

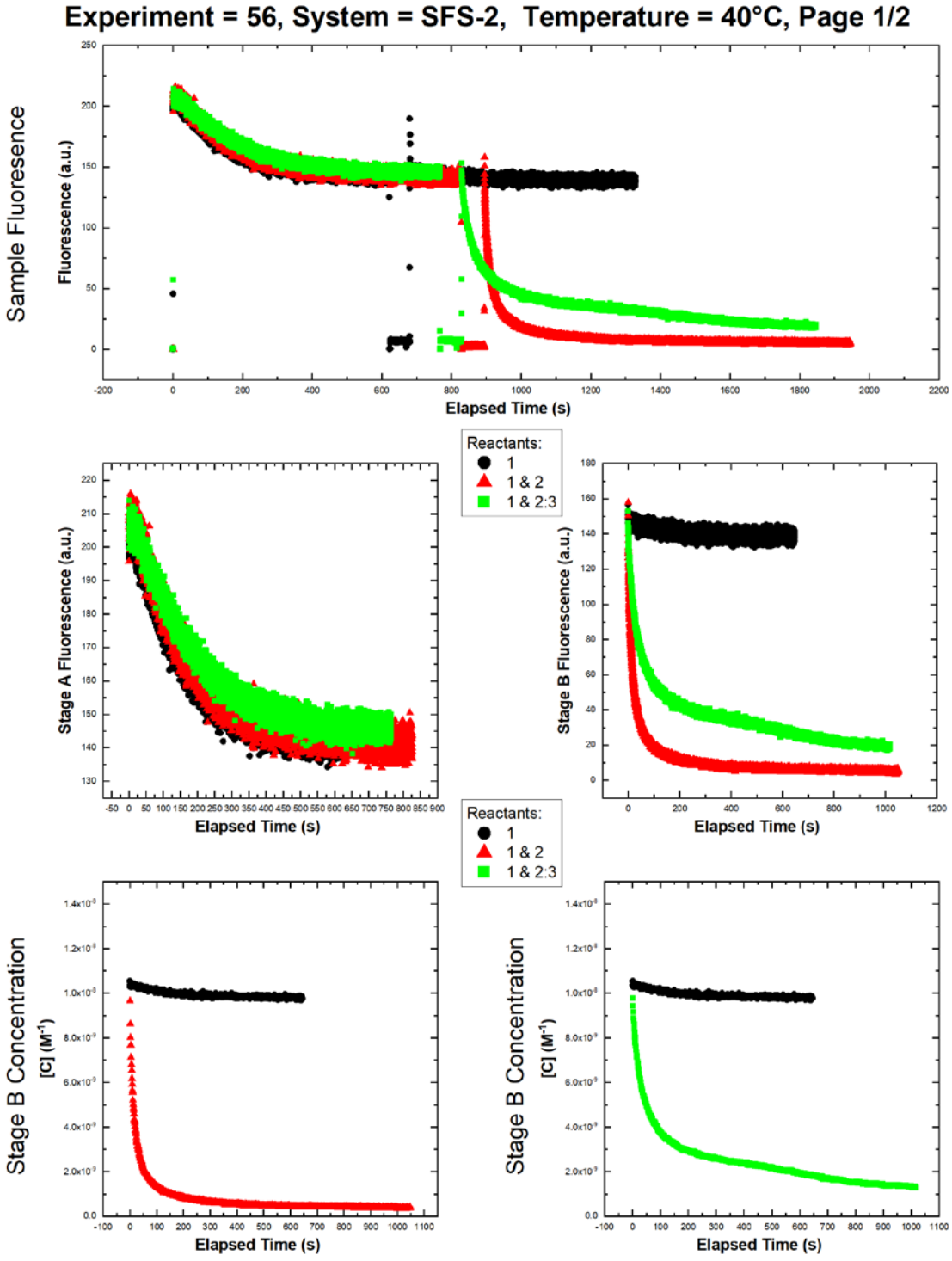

Experiment = 53, System = SFS-2, Temperature = 10°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

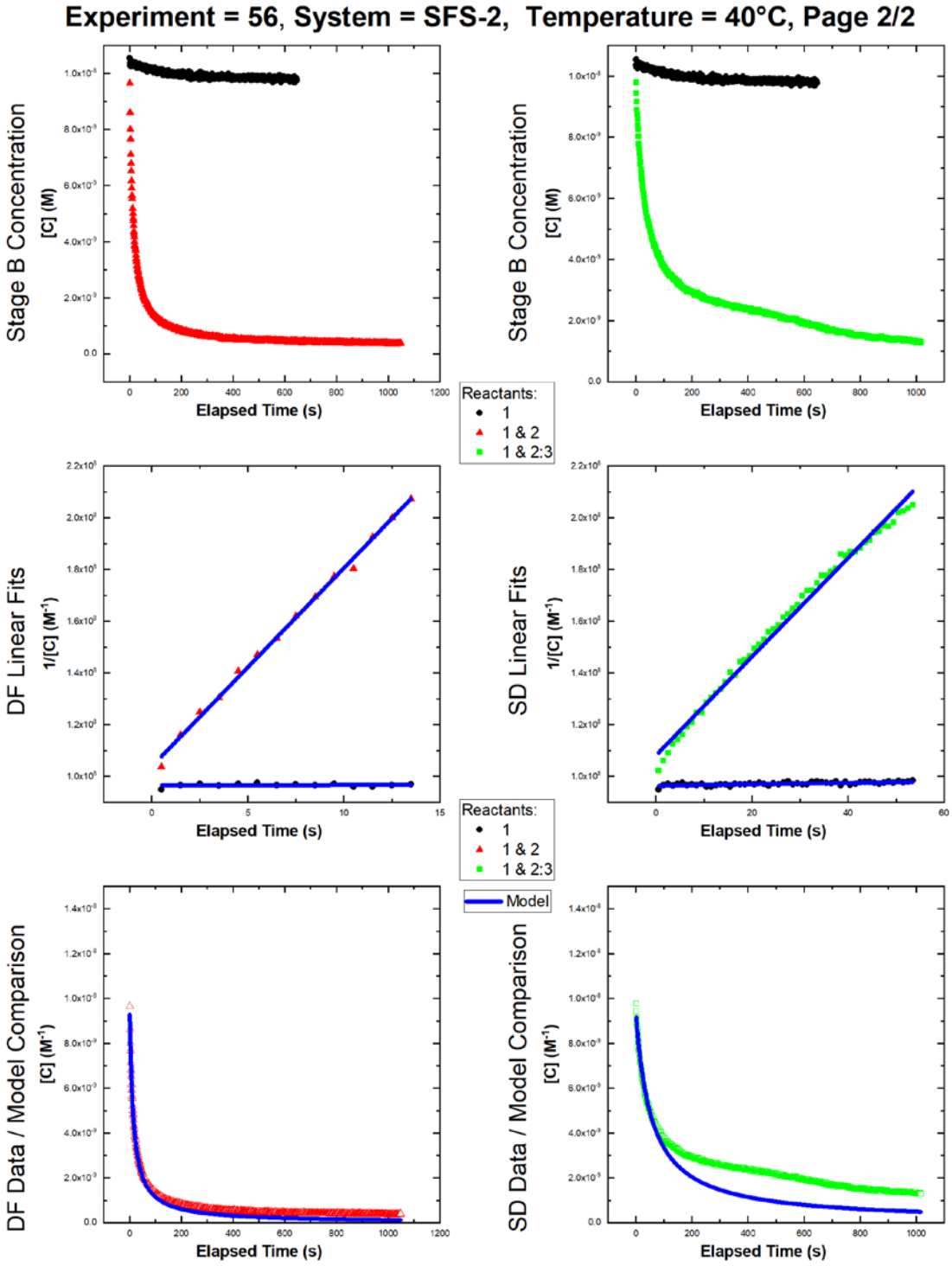

Experiment = 54, System = SFS-2, Temperature = 20°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

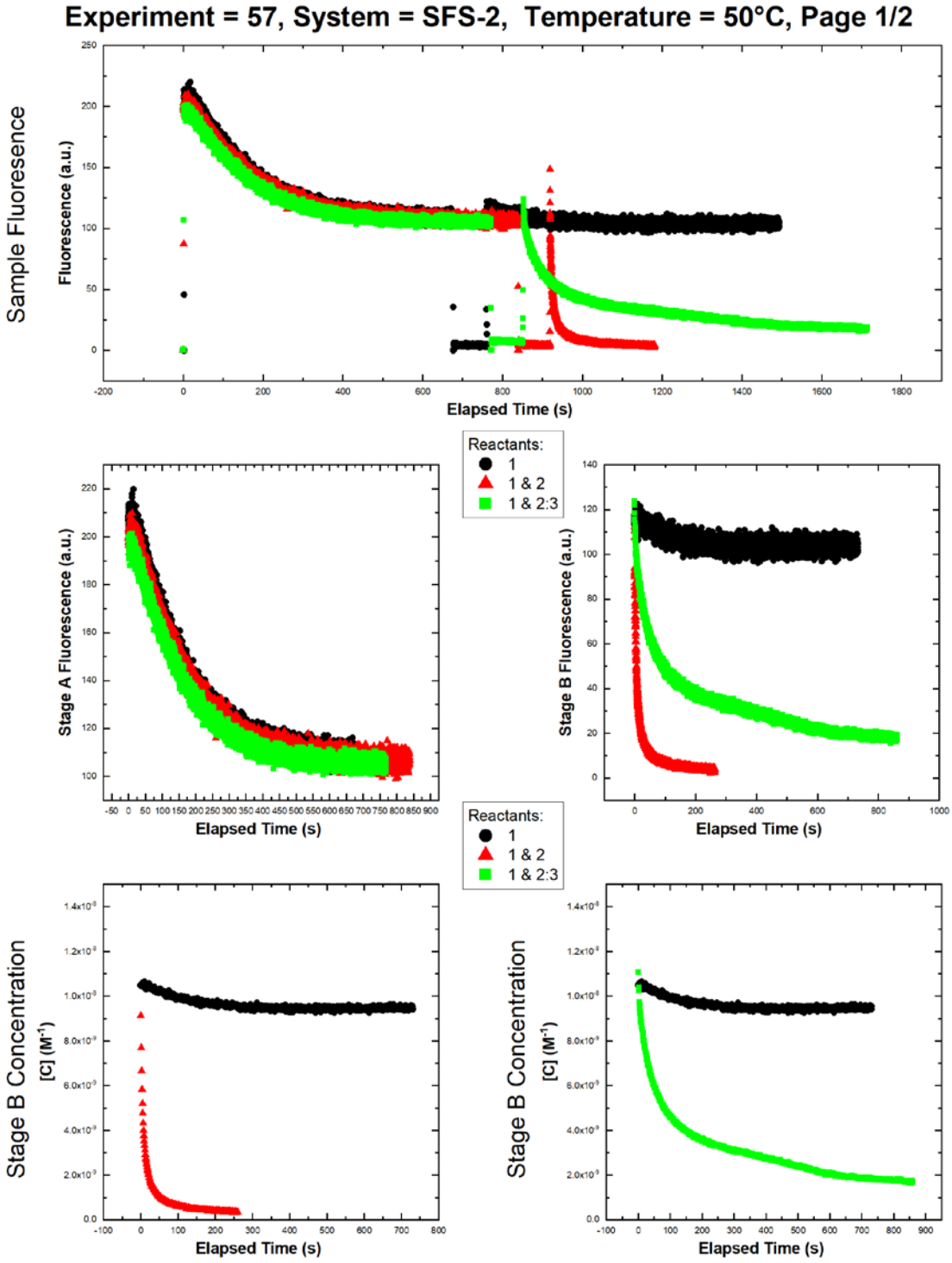

Experiment = 54, System = SFS-2, Temperature = 20°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

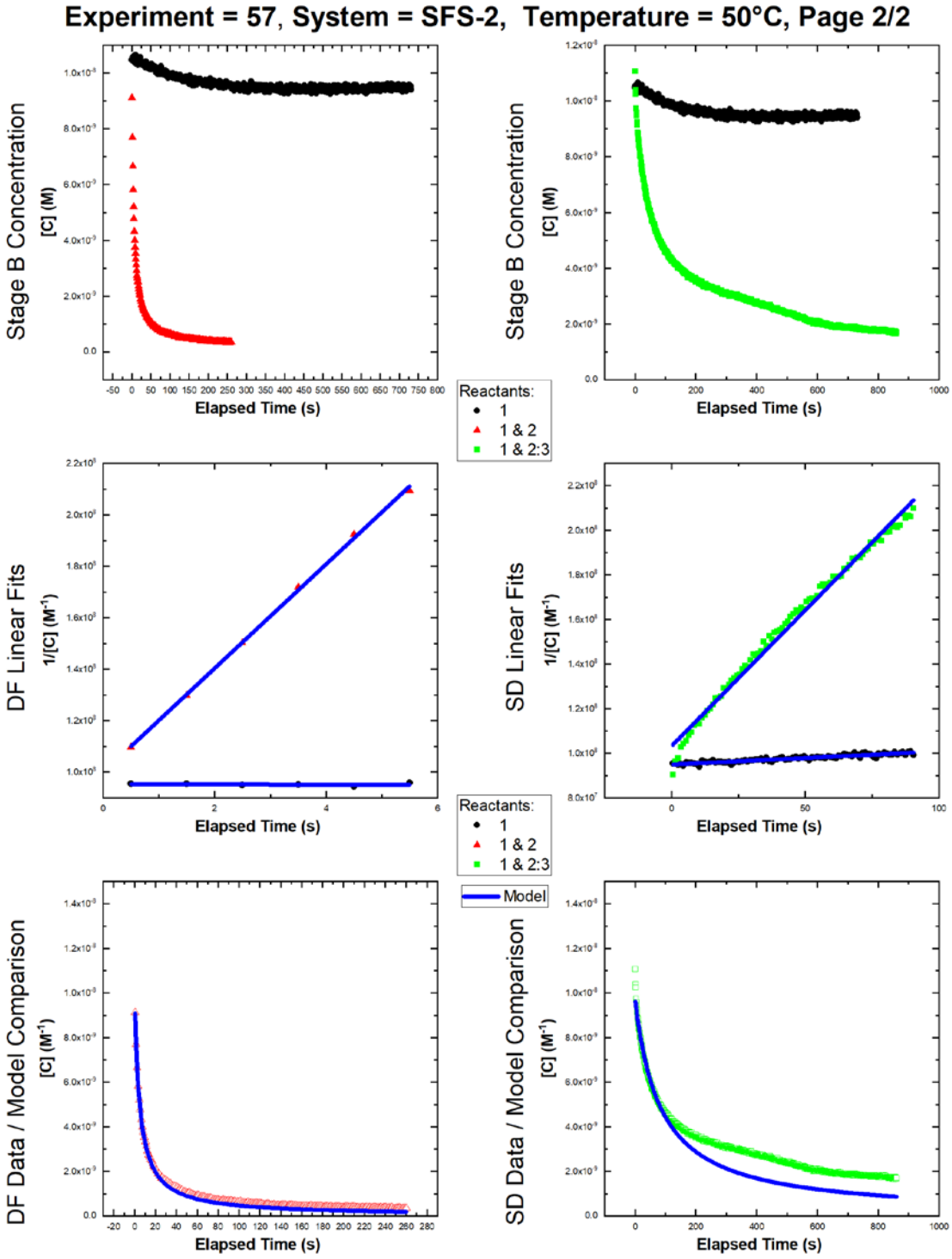

Experiment = 55, System = SFS-2, Temperature = 30°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

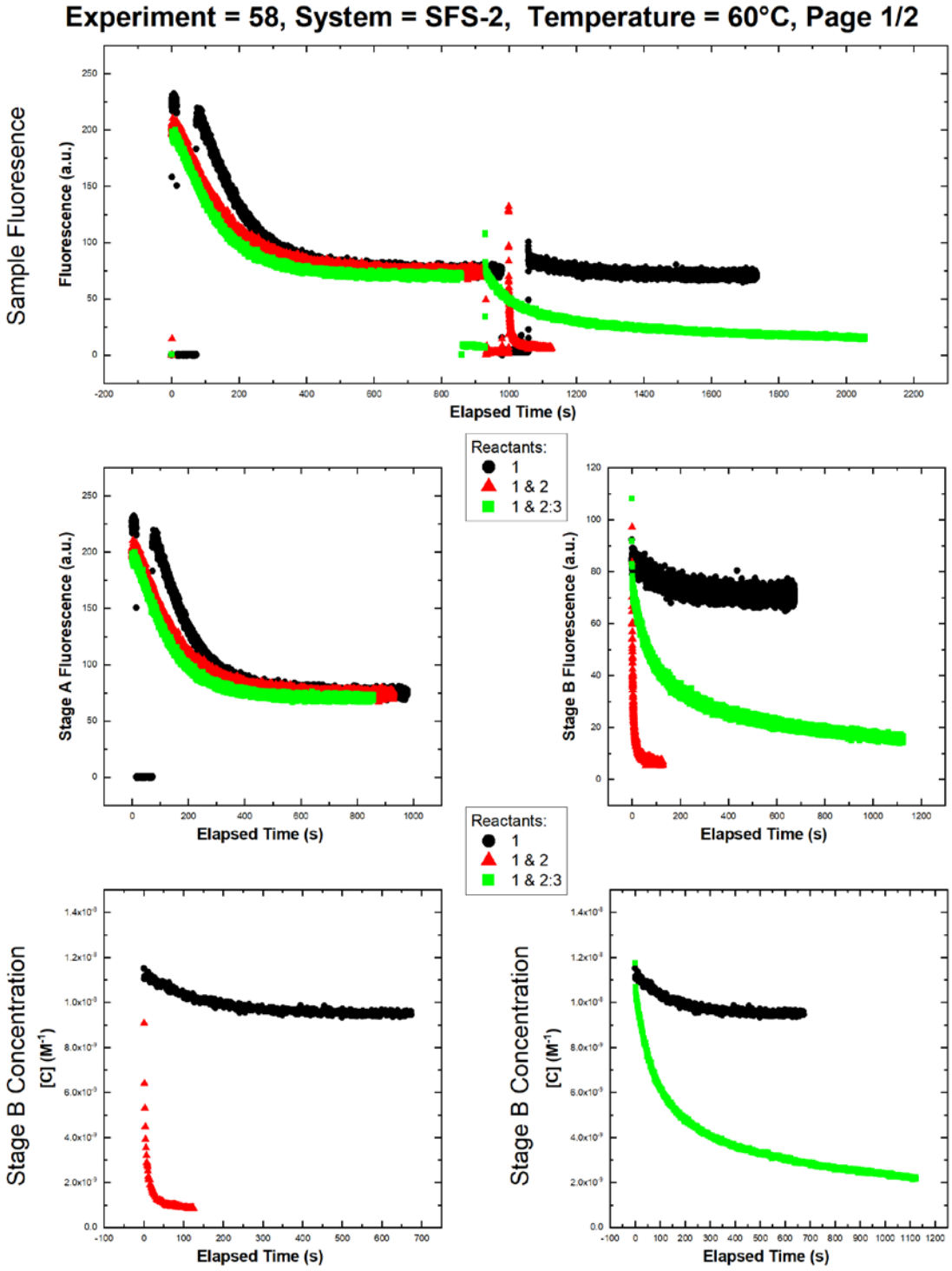

Experiment = 55, System = SFS-2, Temperature = 30°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

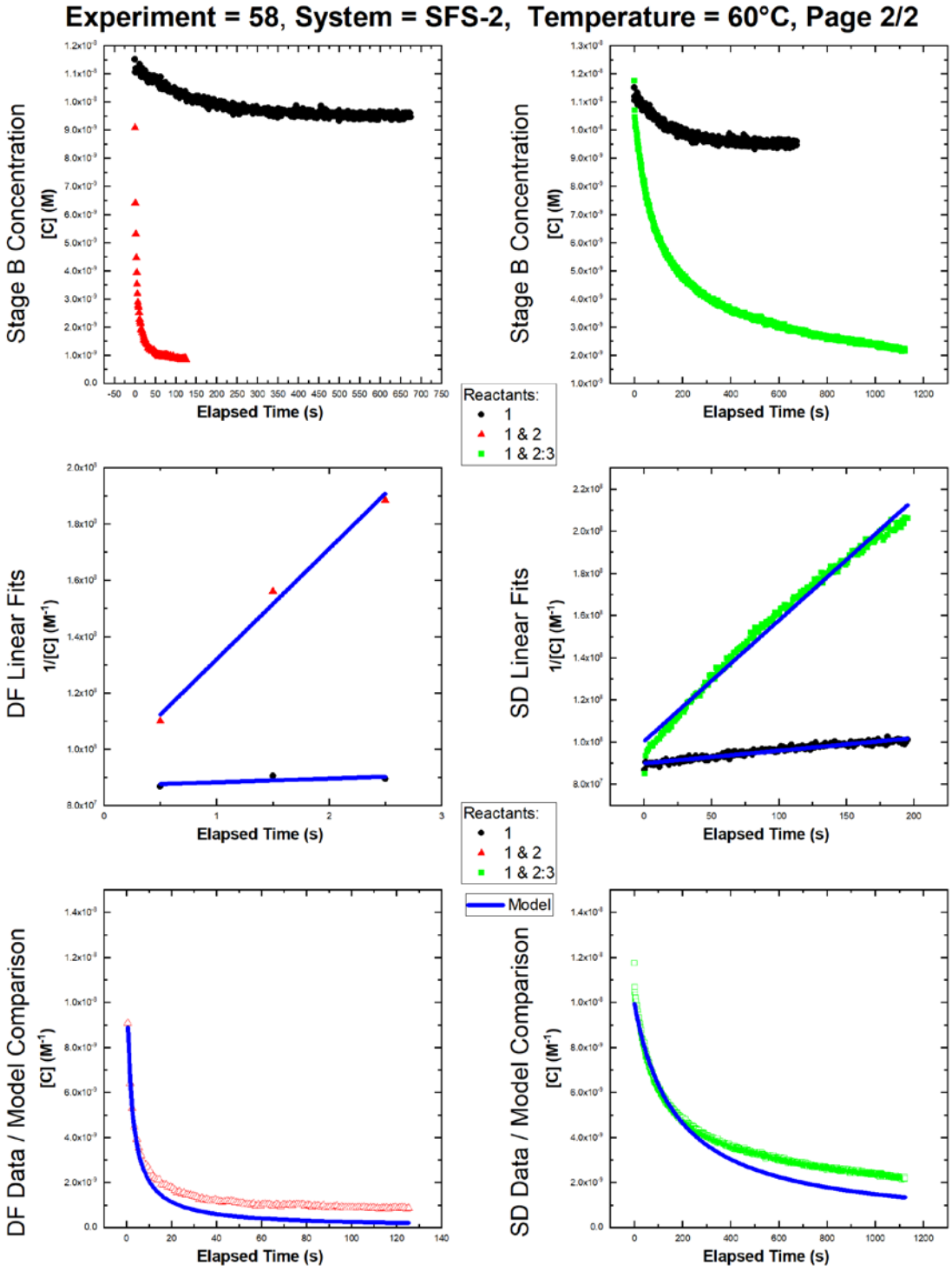

Experiment = 56, System = SFS-2, Temperature = 40°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

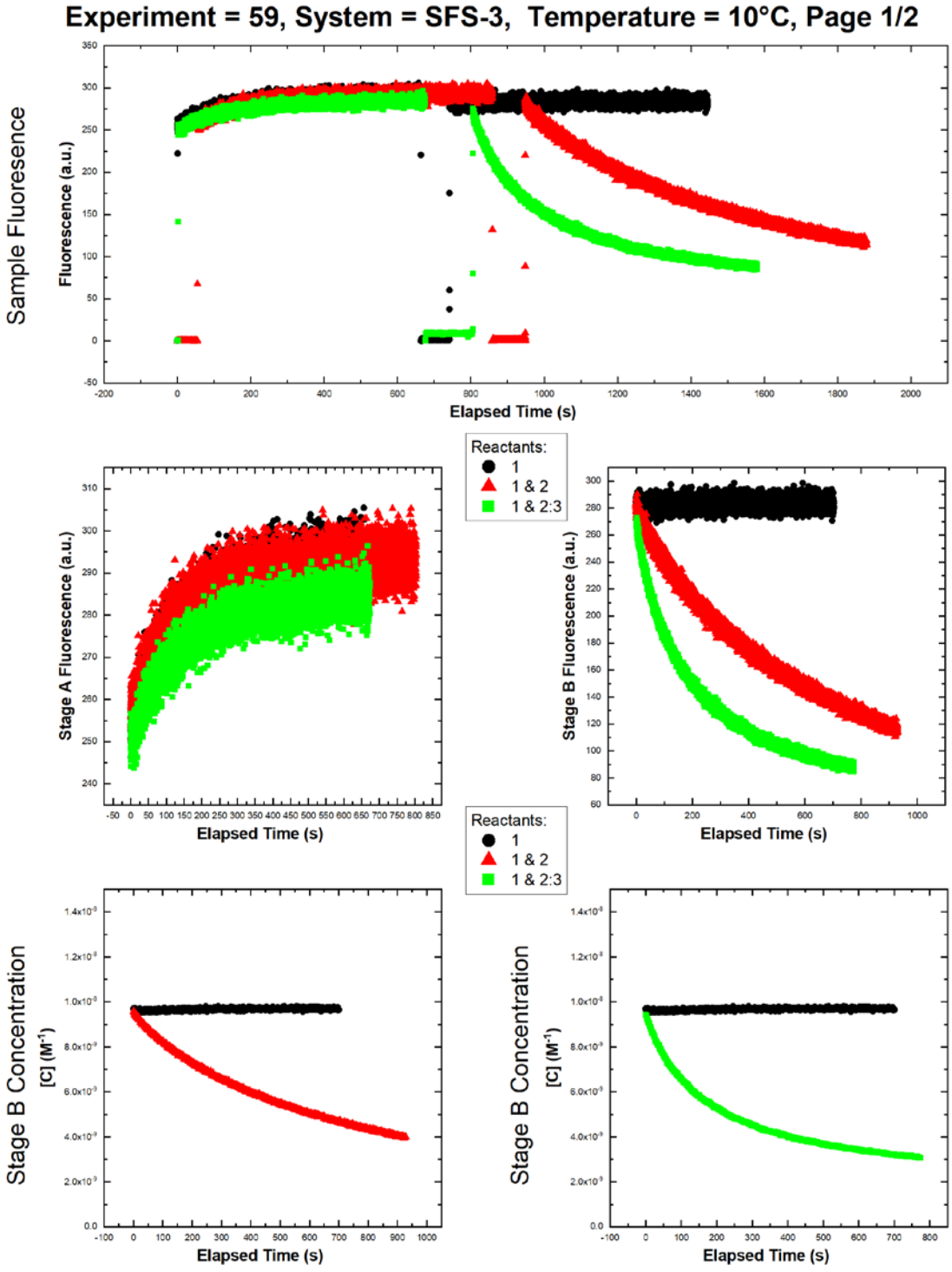

Experiment = 56, System = SFS-2, Temperature = 40°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

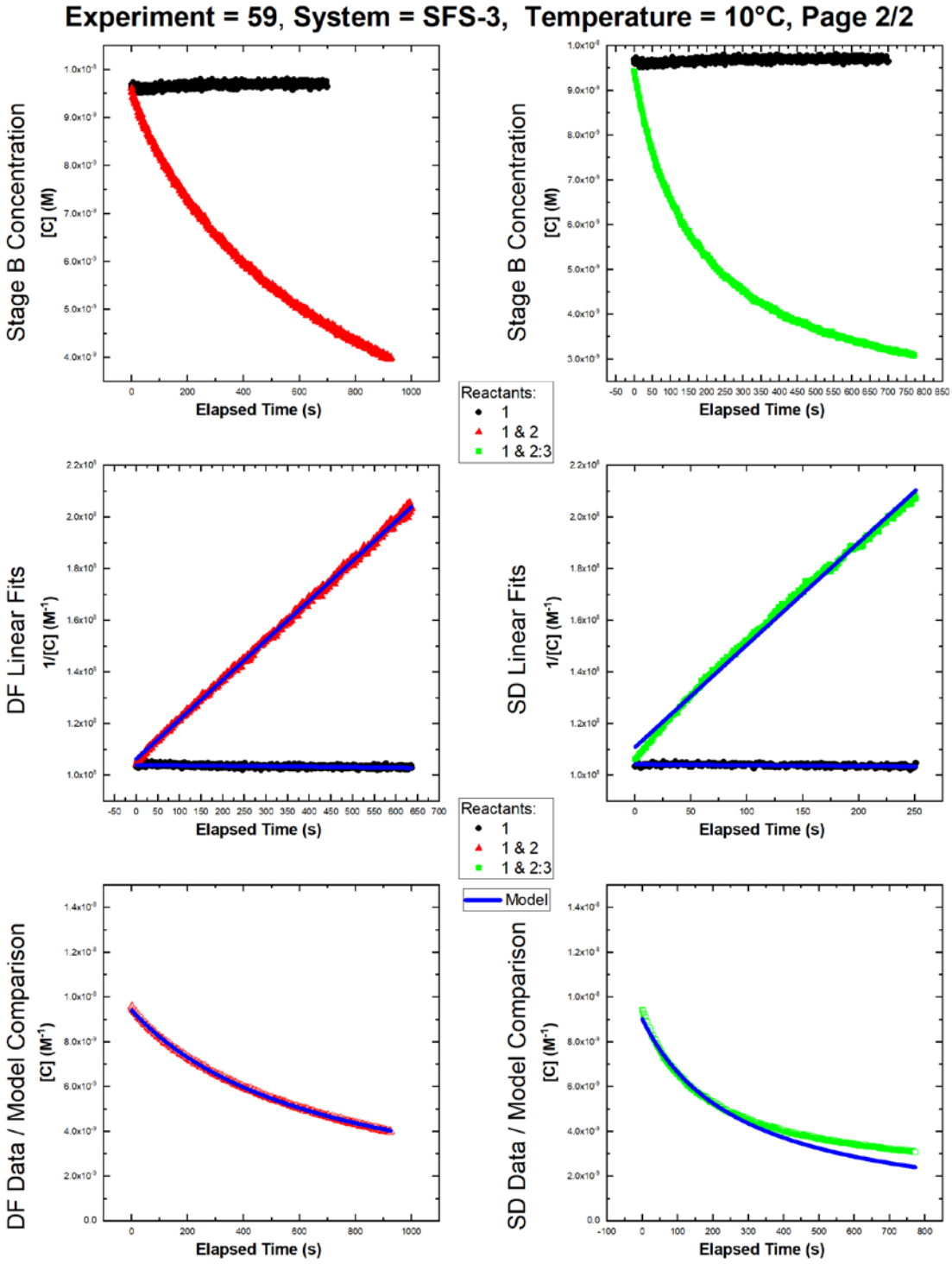

Experiment = 57, System = SFS-2, Temperature = 50°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

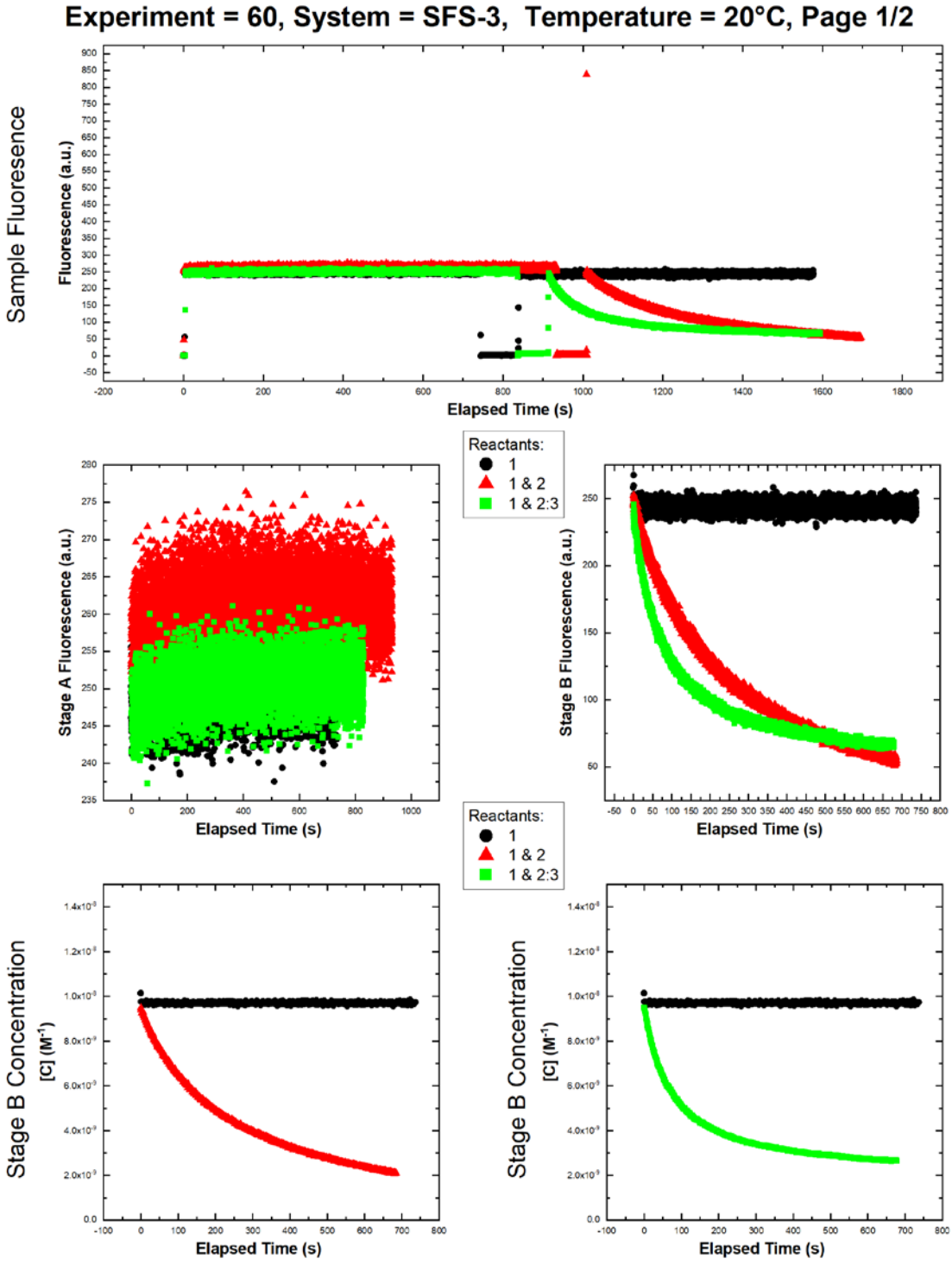

Experiment = 57, System = SFS-2, Temperature = 50°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

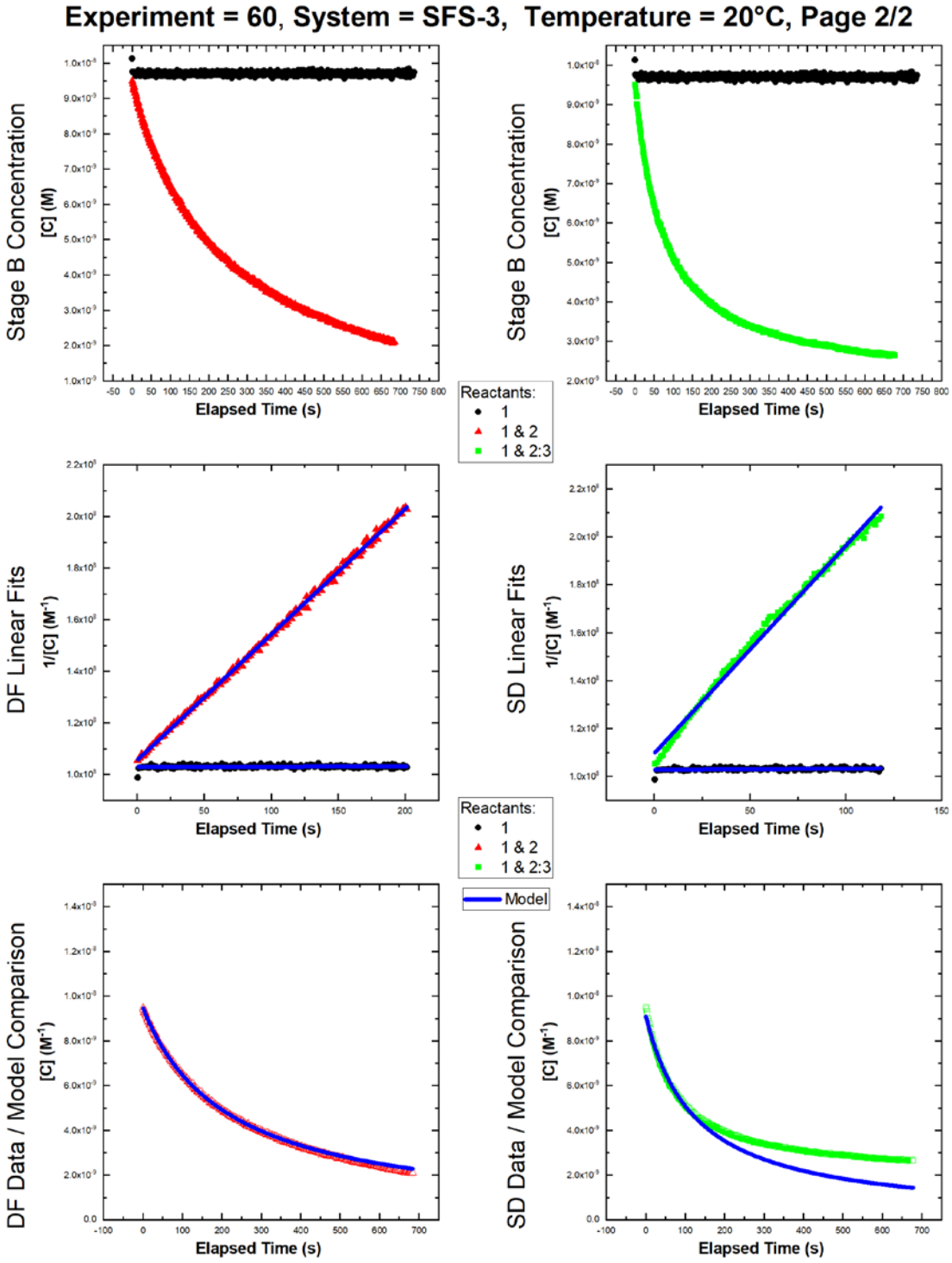

Experiment = 58, System = SFS-2, Temperature = 60°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

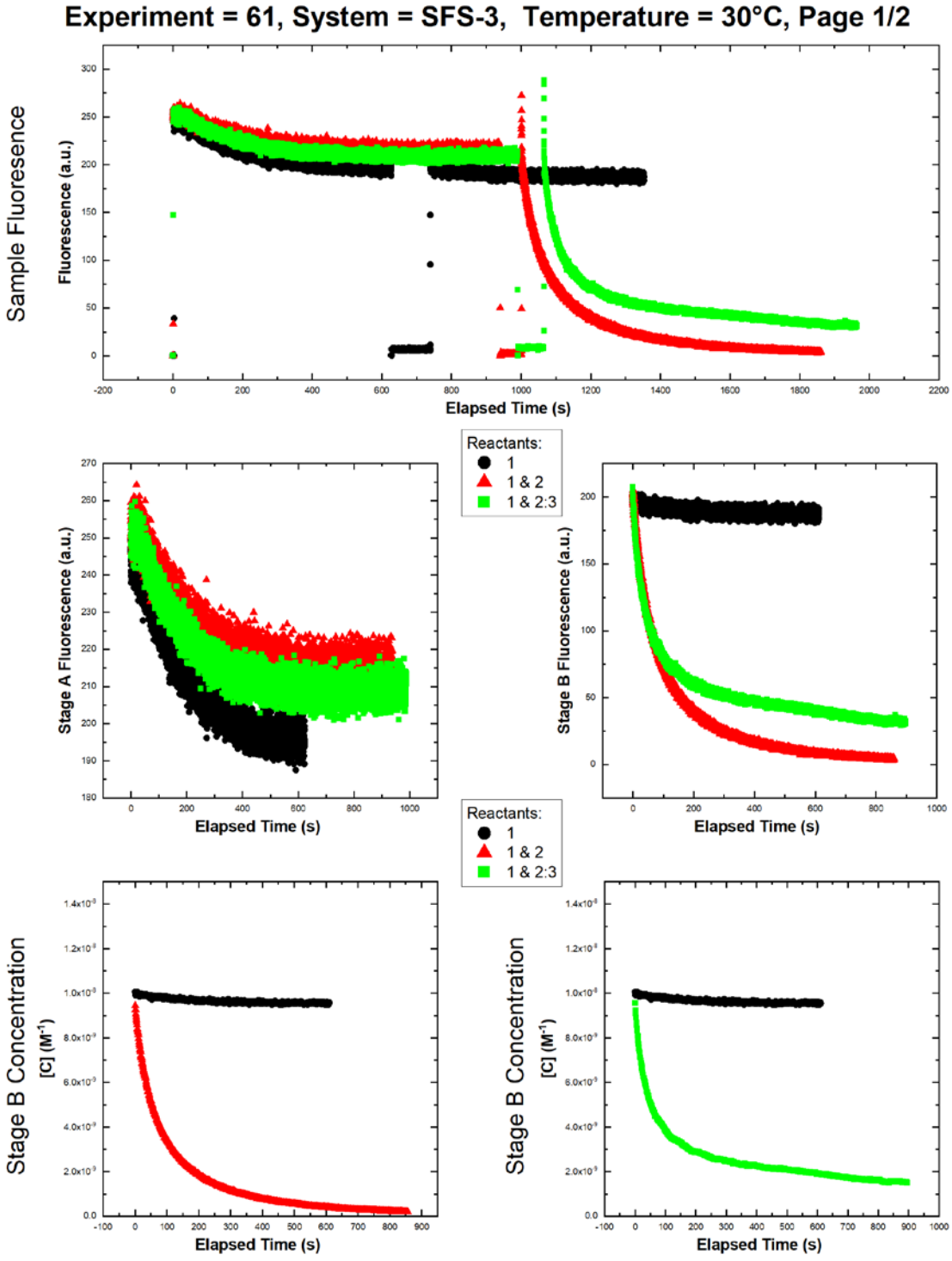

Experiment = 58, System = SFS-2, Temperature = 60°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

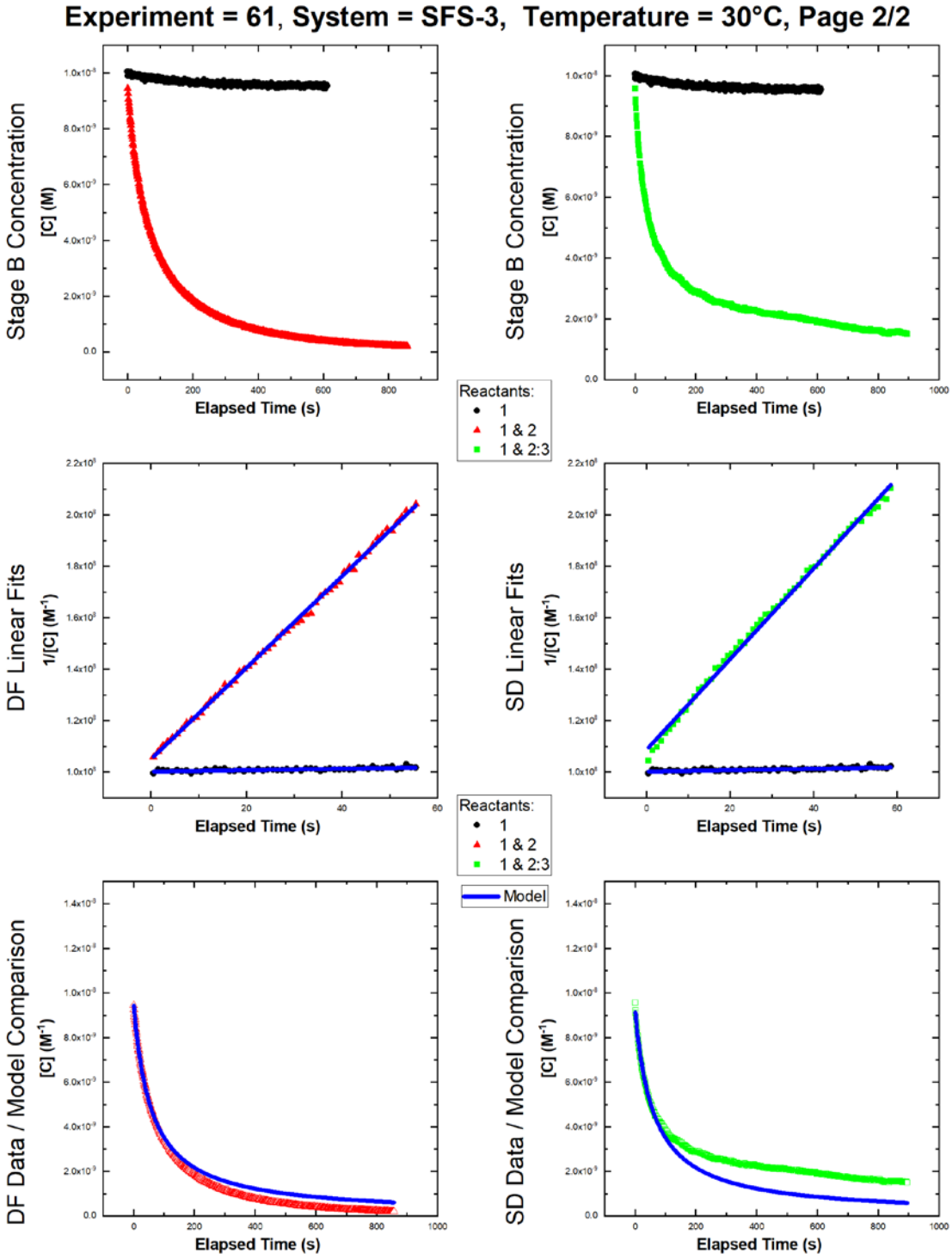

Experiment = 59, System = SFS-3, Temperature = 10°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

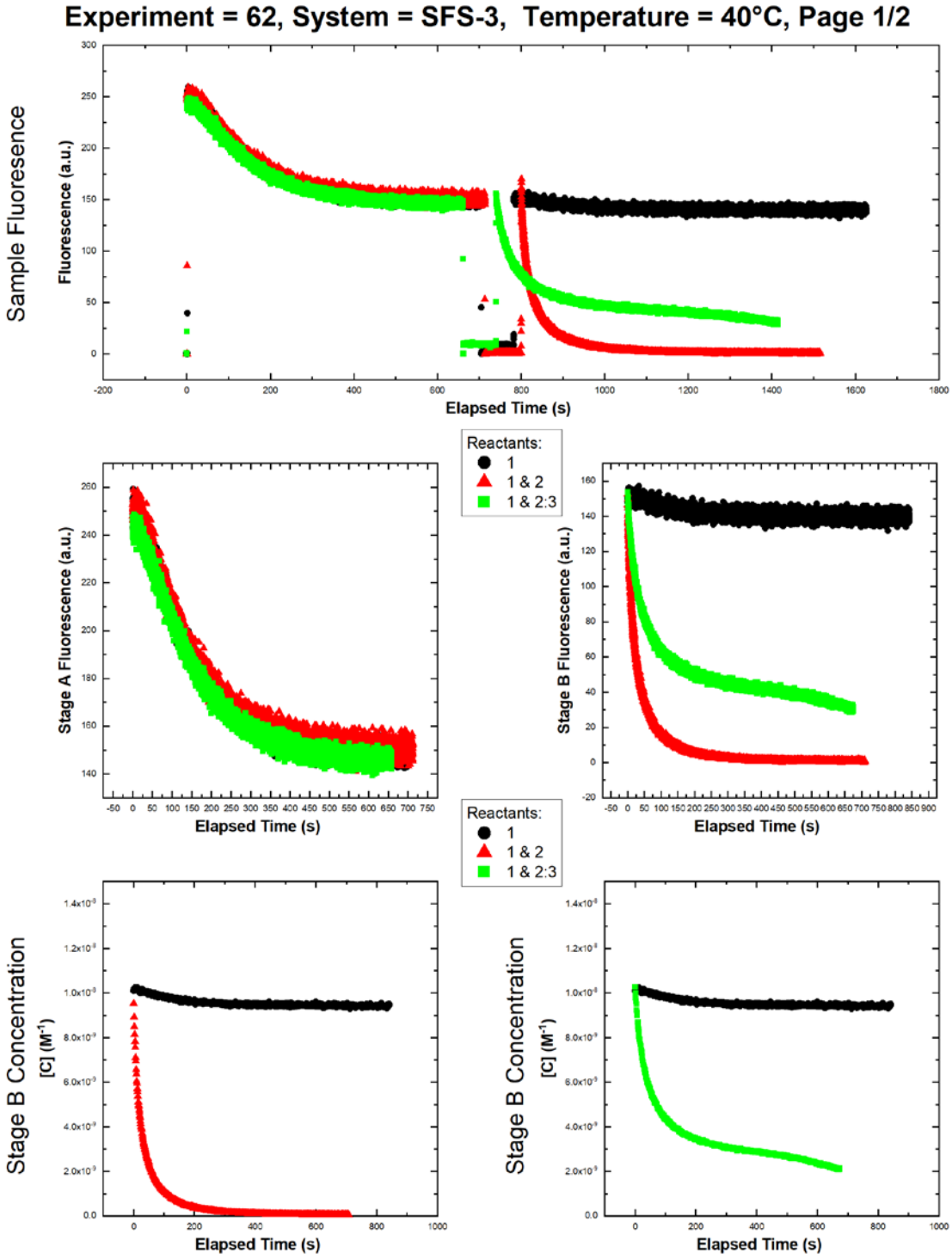

Experiment = 59, System = SFS-3, Temperature = 10°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

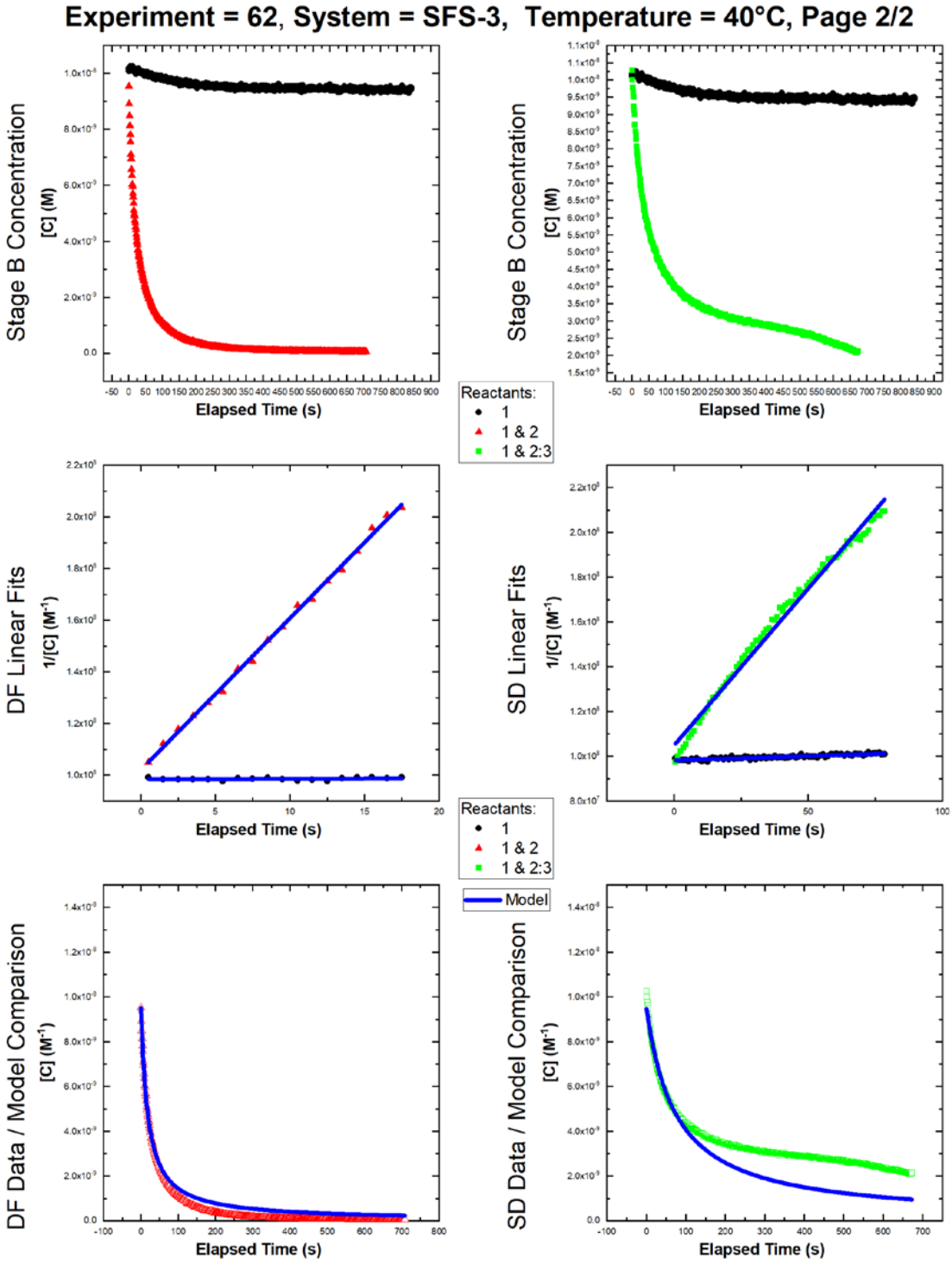

Experiment = 60, System = SFS-3, Temperature = 20°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

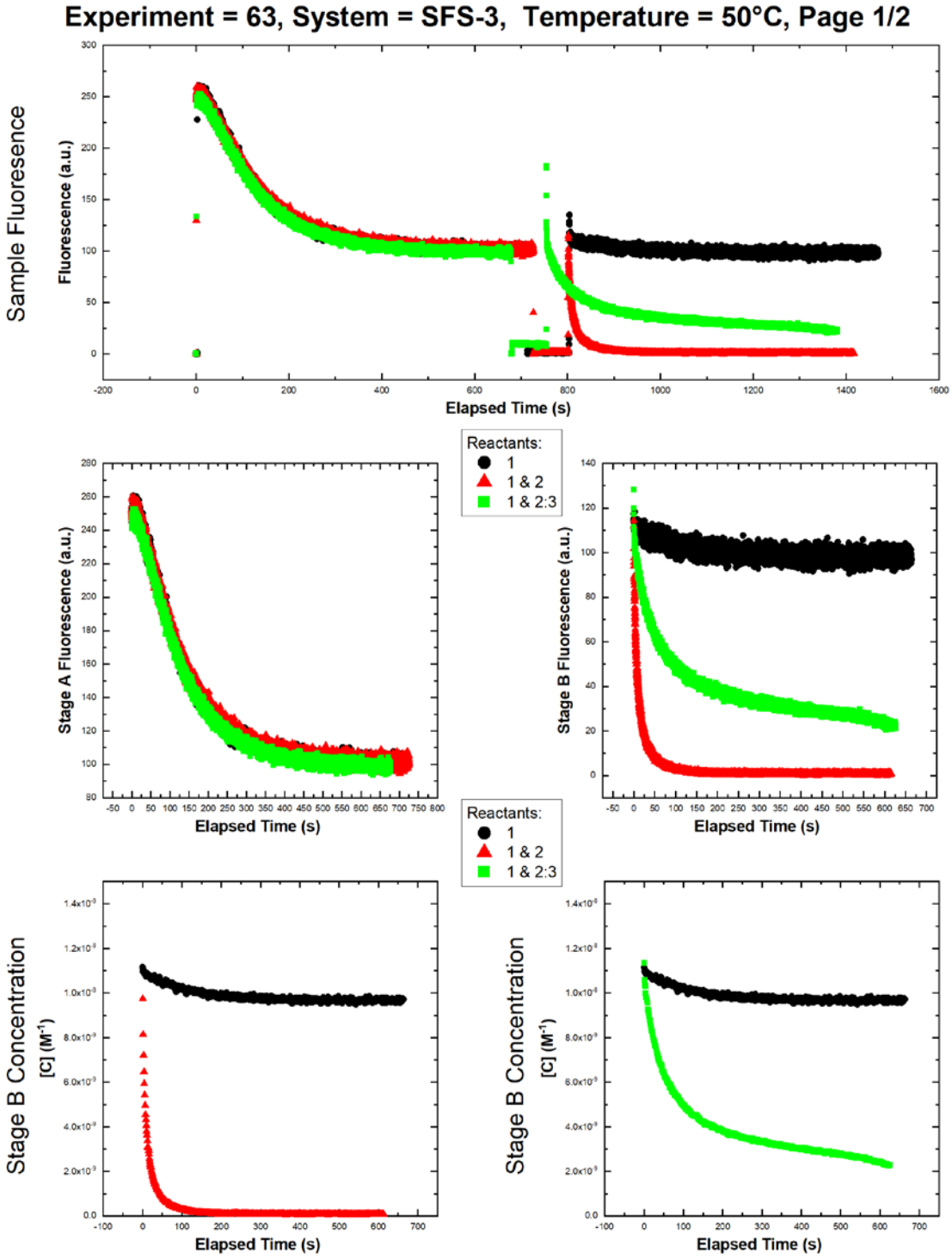

Experiment = 60, System = SFS-3, Temperature = 20°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

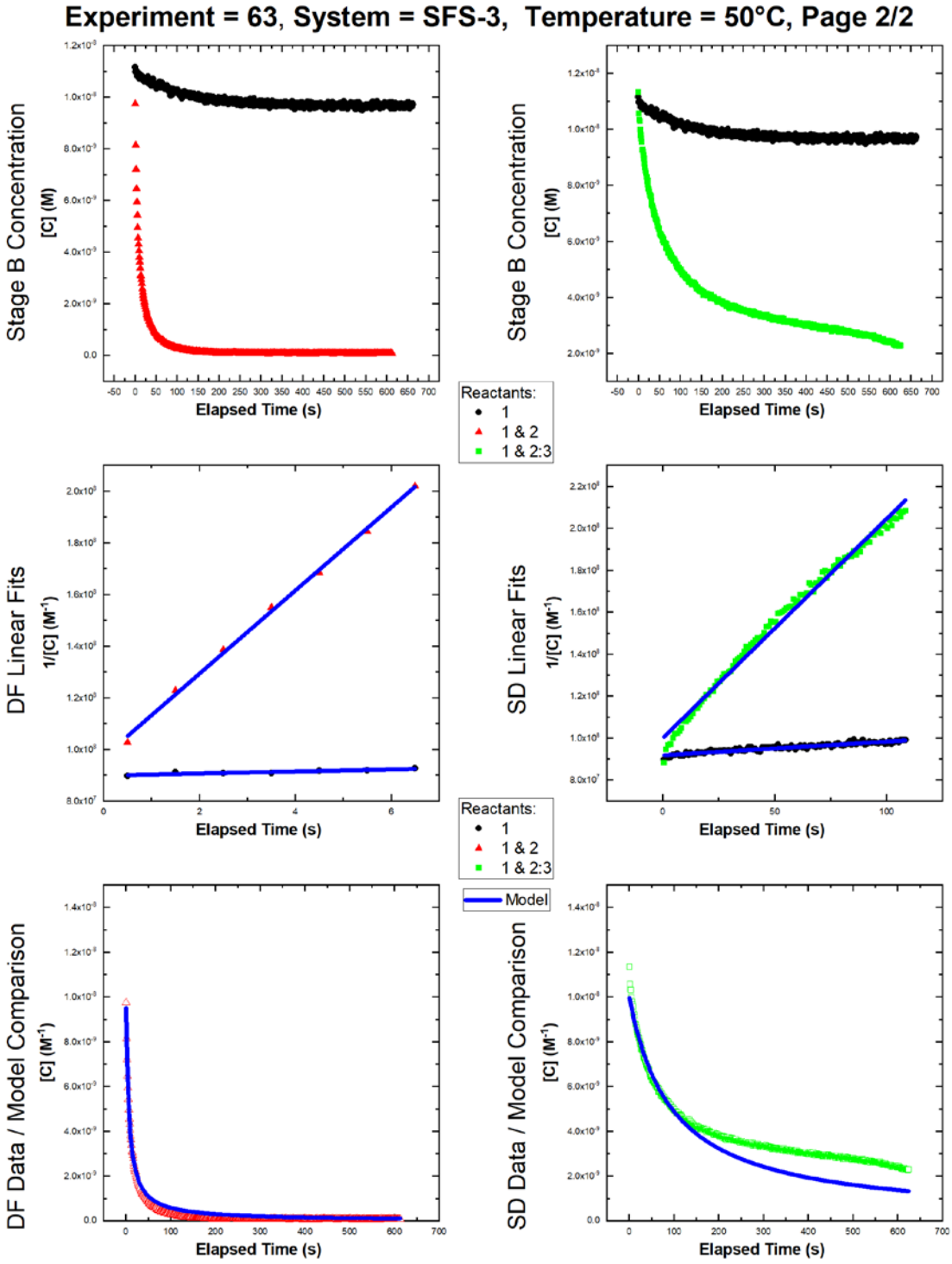

Experiment = 61, System = SFS-3, Temperature = 30°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

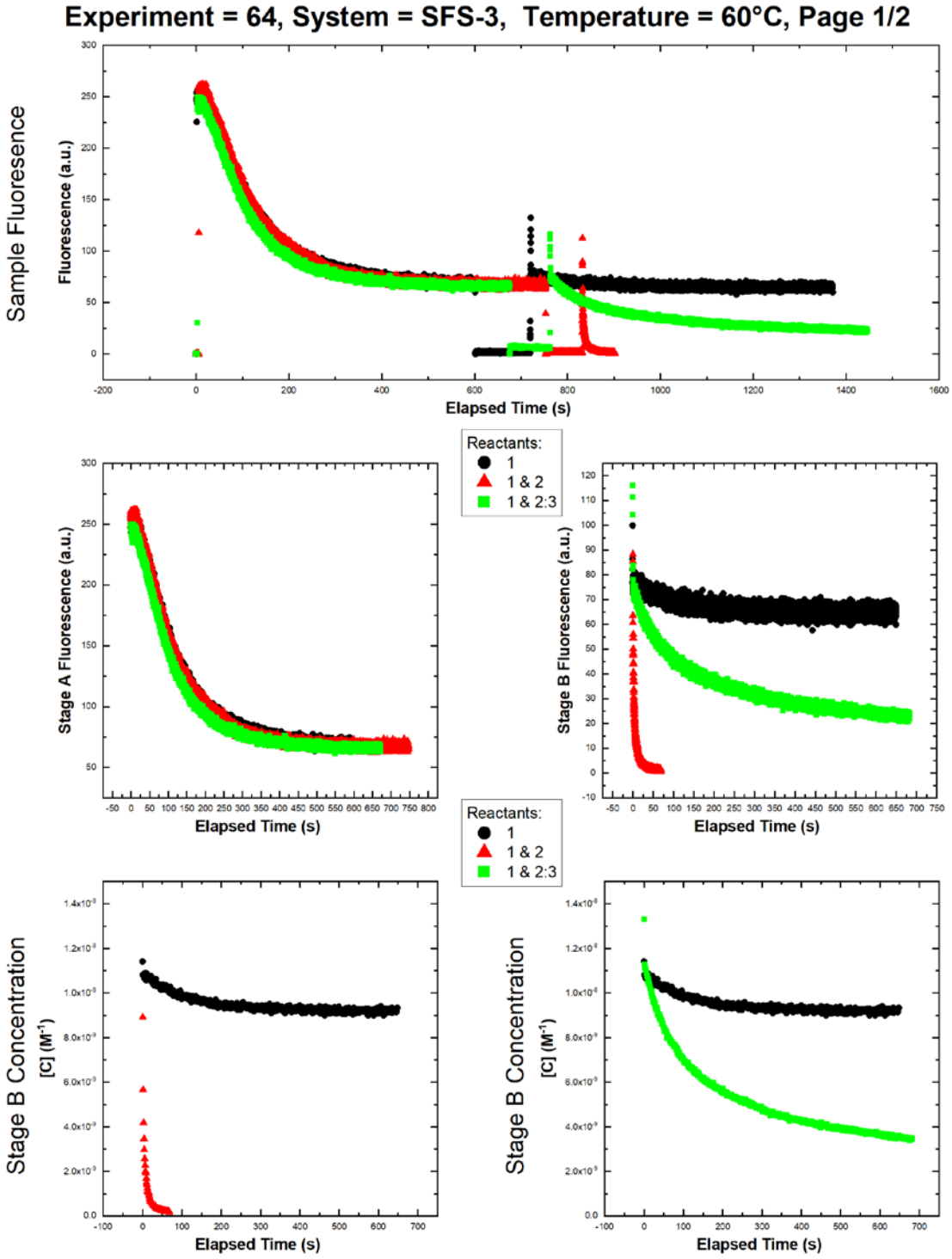

Experiment = 61, System = SFS-3, Temperature = 30°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
[C] (M-1)
SD Data / Model Comparison

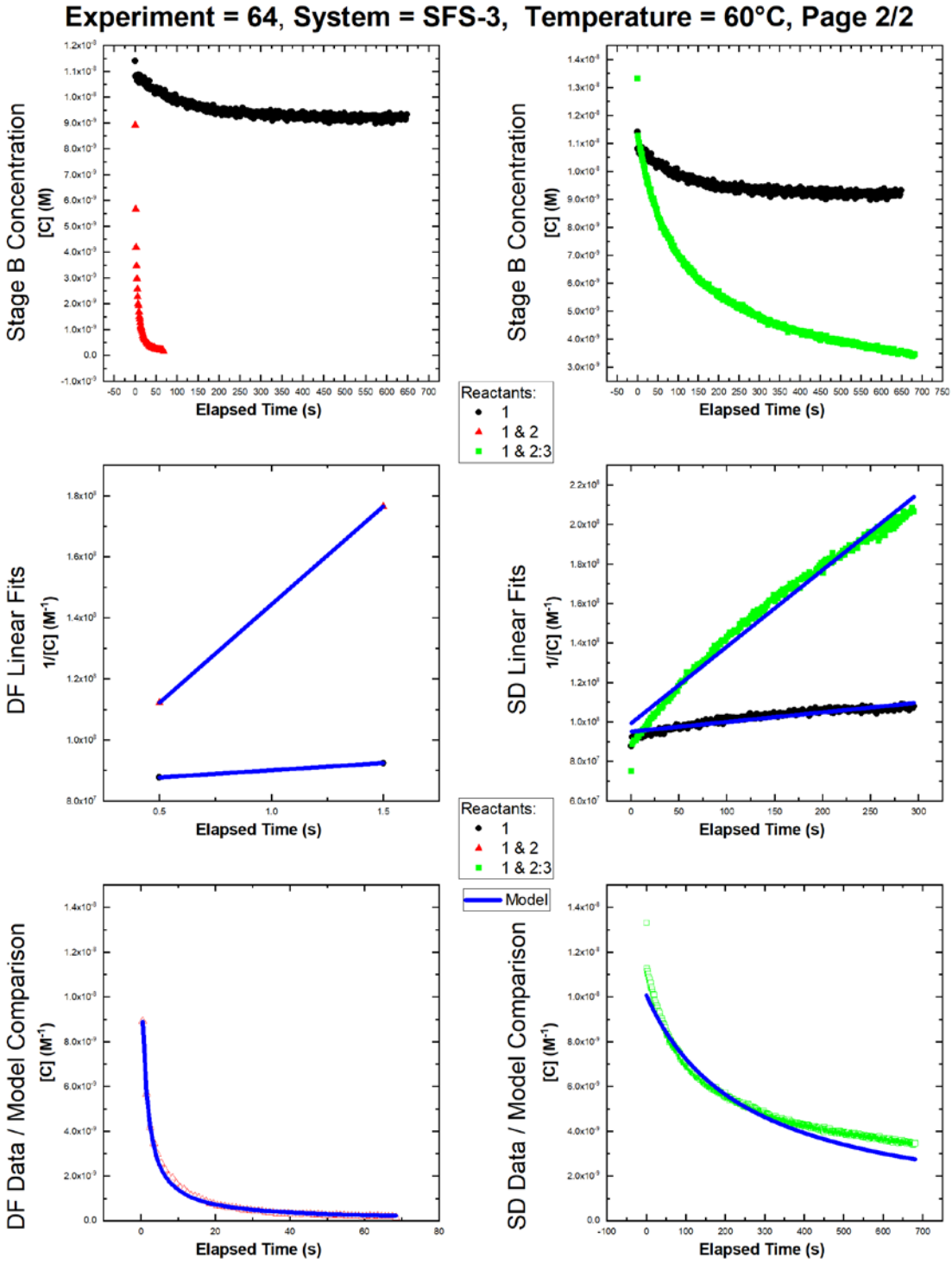

Experiment = 62, System = SFS-3, Temperature = 40°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

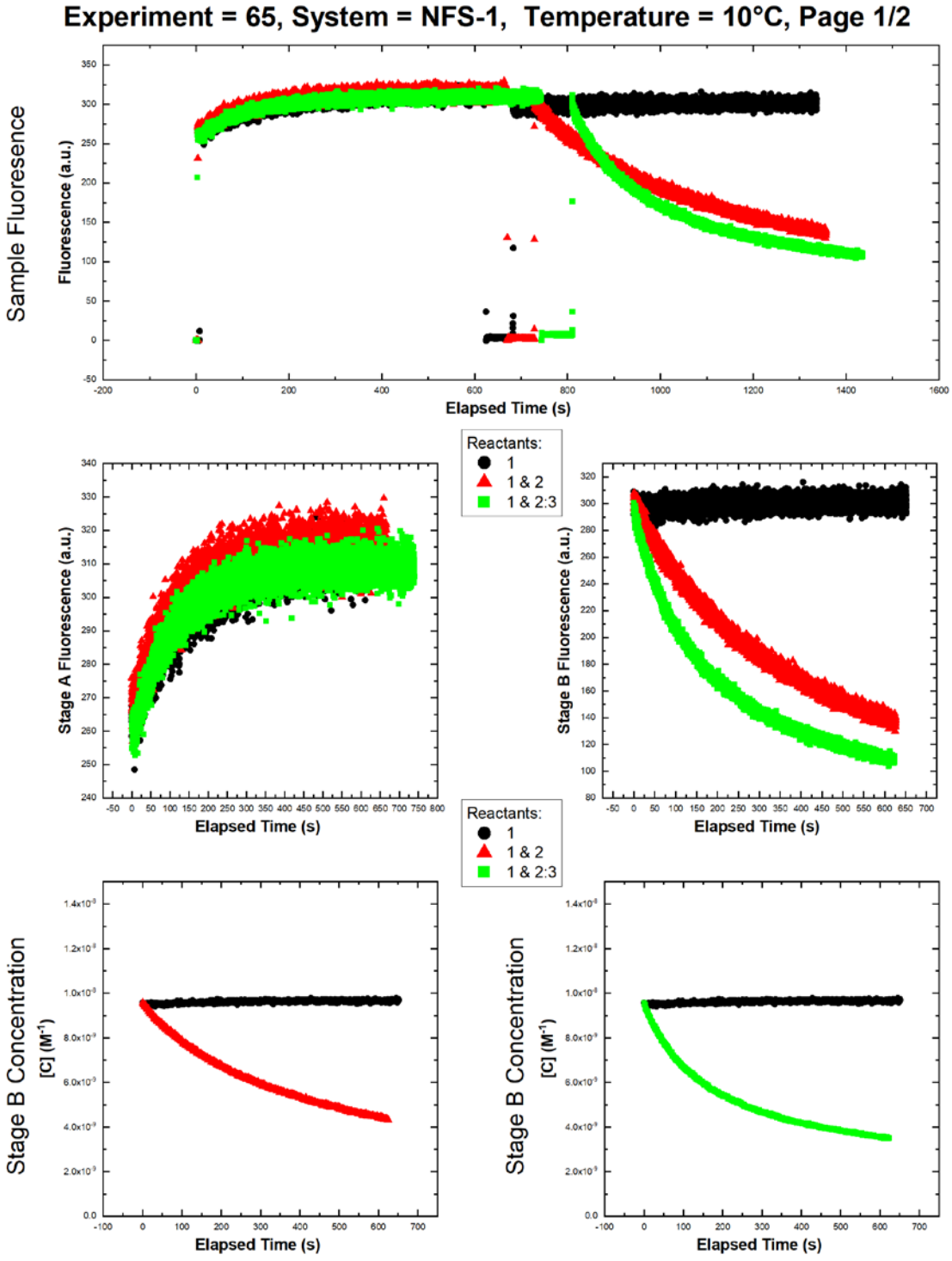

Experiment = 62, System = SFS-3, Temperature = 40°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

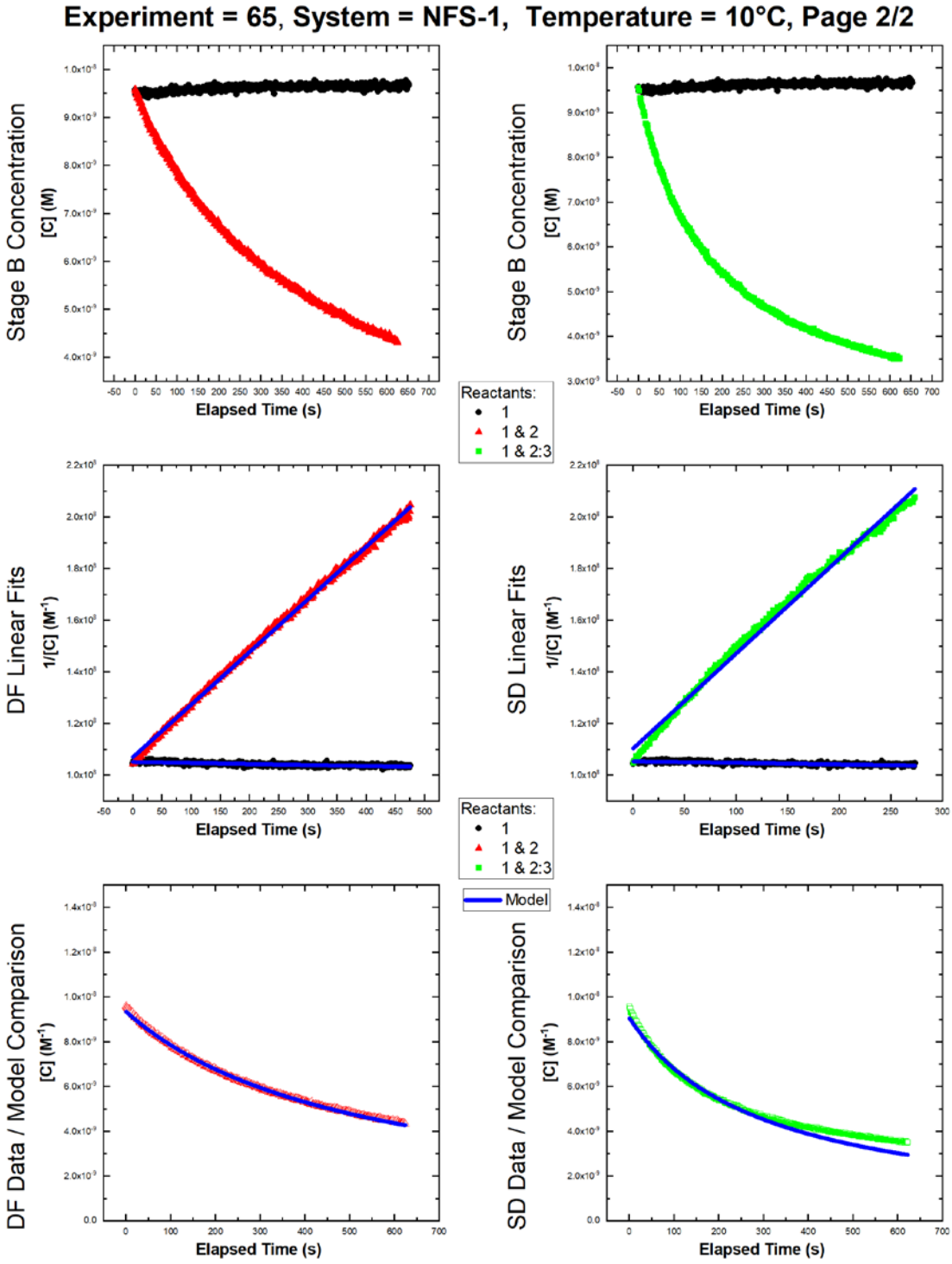

Experiment = 63, System = SFS-3, Temperature = 50°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

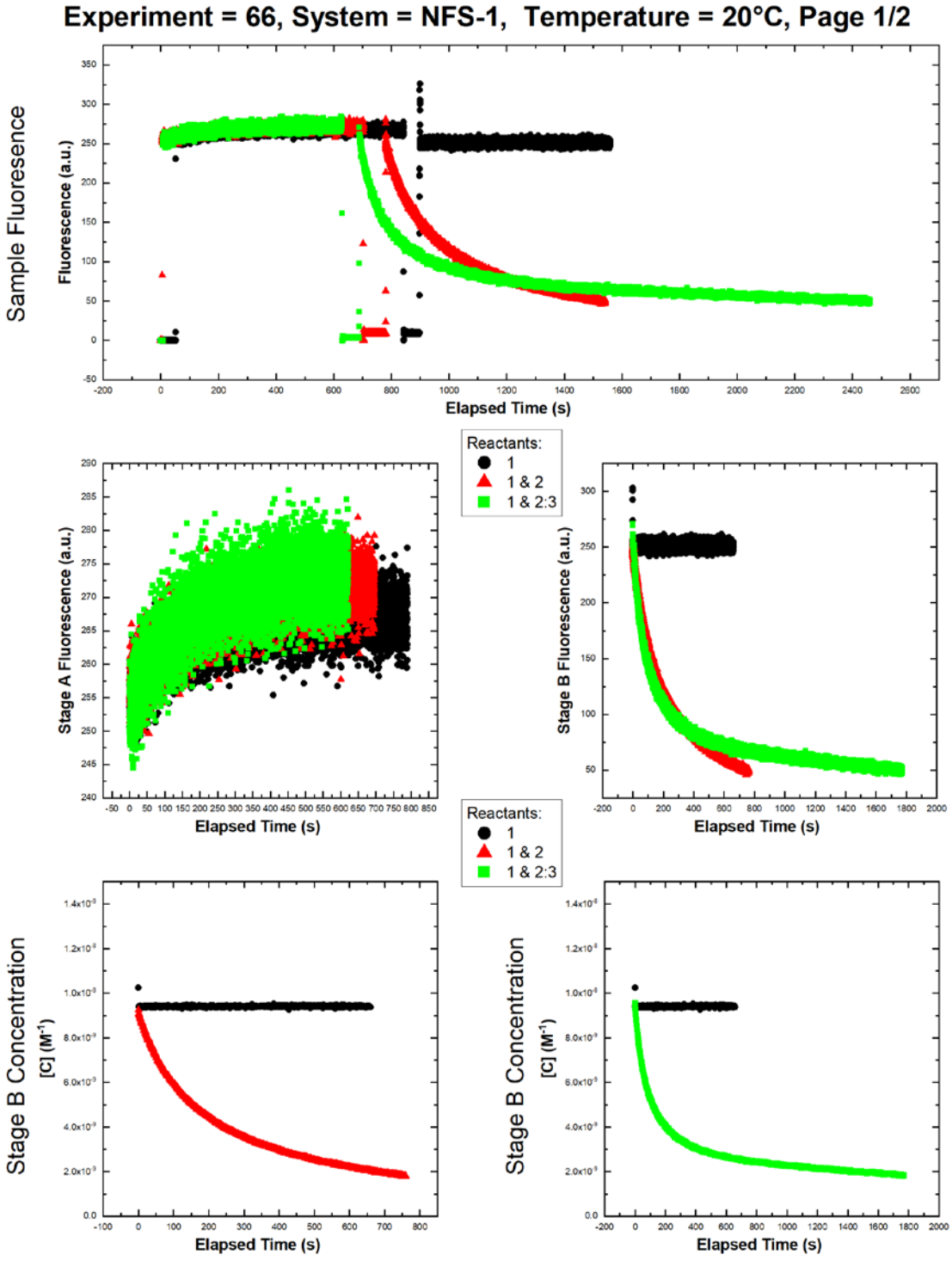

Experiment = 63, System = SFS-3, Temperature = 50°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

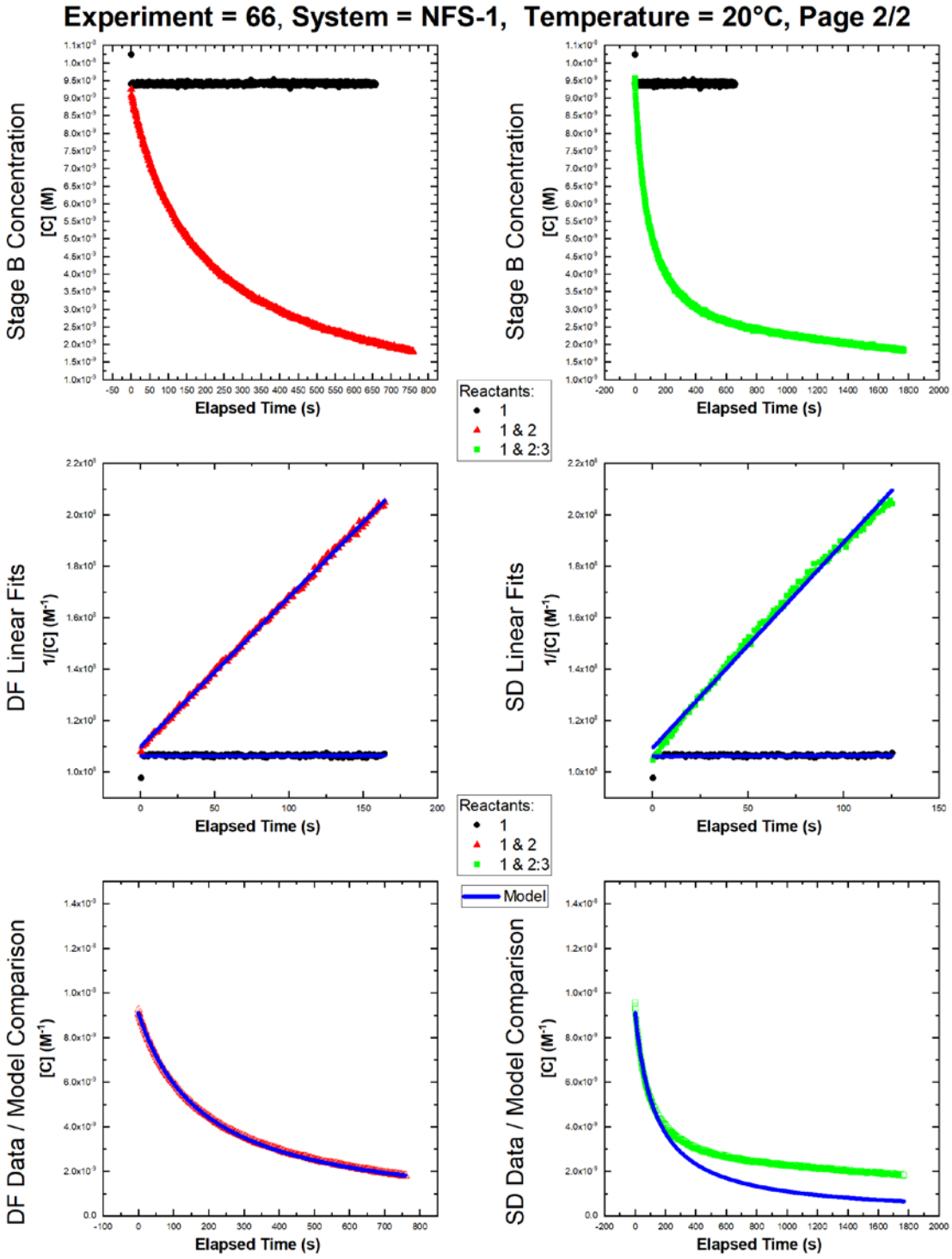

Experiment = 64, System = SFS-3, Temperature = 60°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

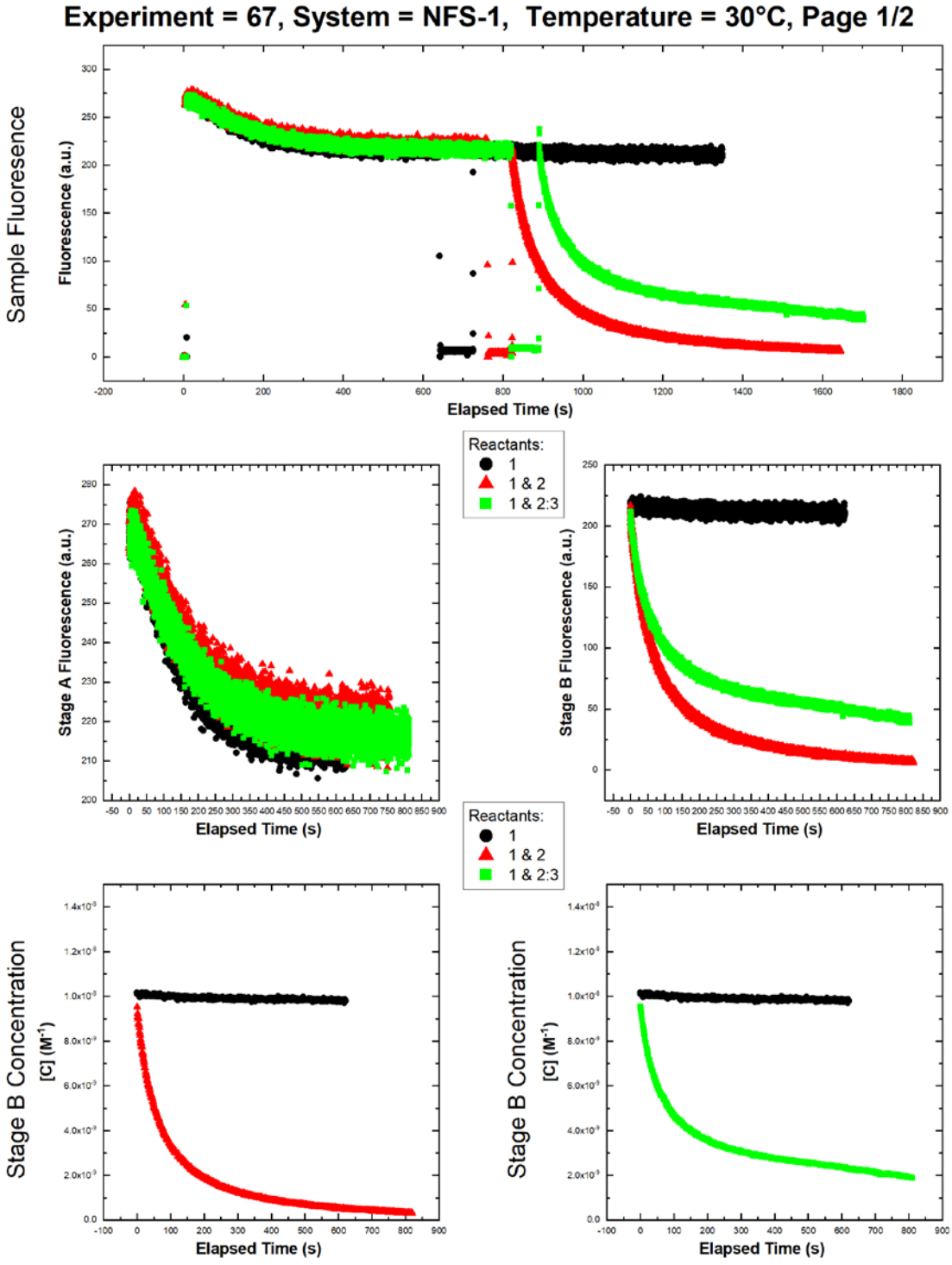

Experiment = 64, System = SFS-3, Temperature = 60°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

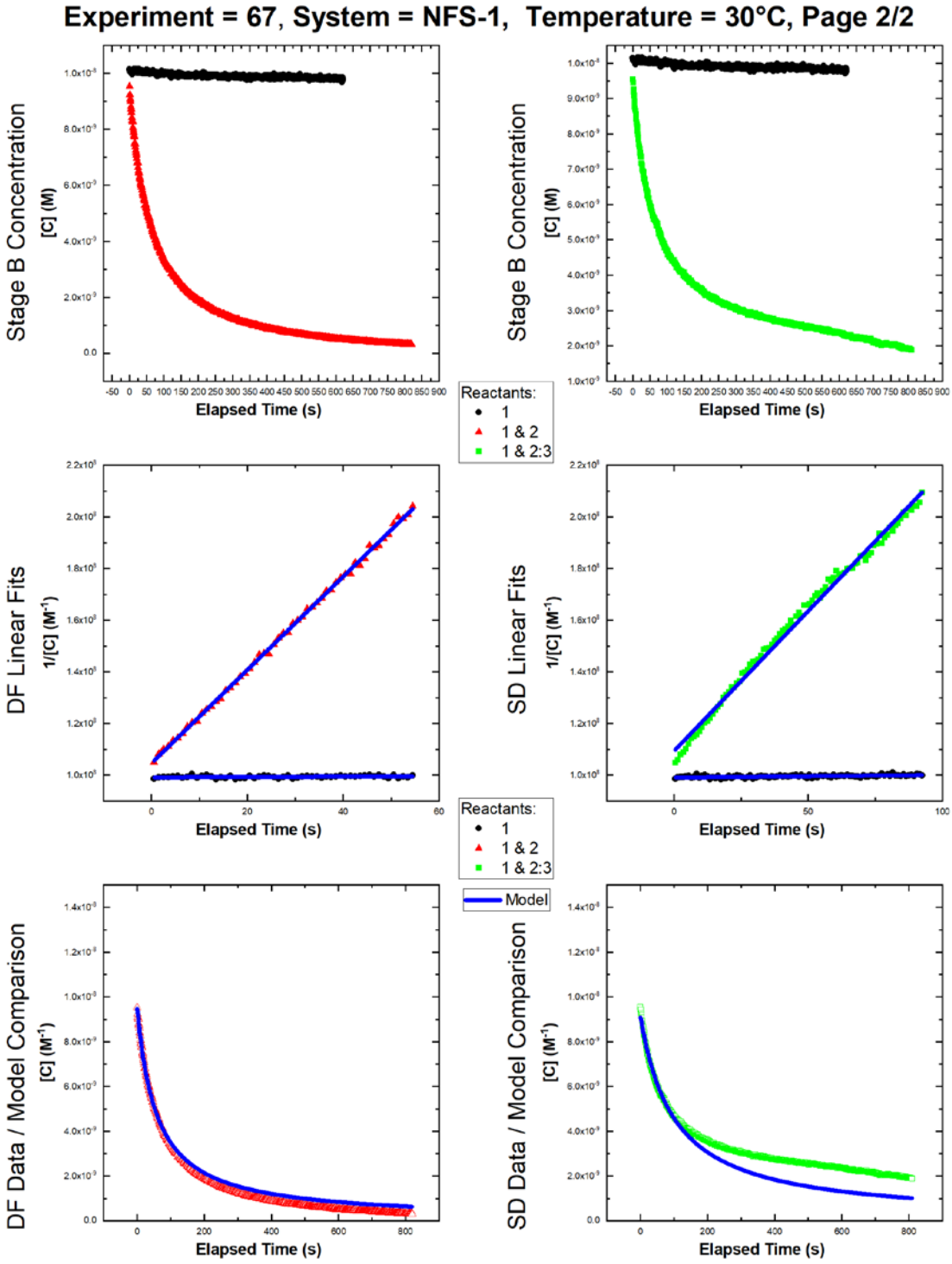

Experiment = 65, System = NFS-1, Temperature = 10°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

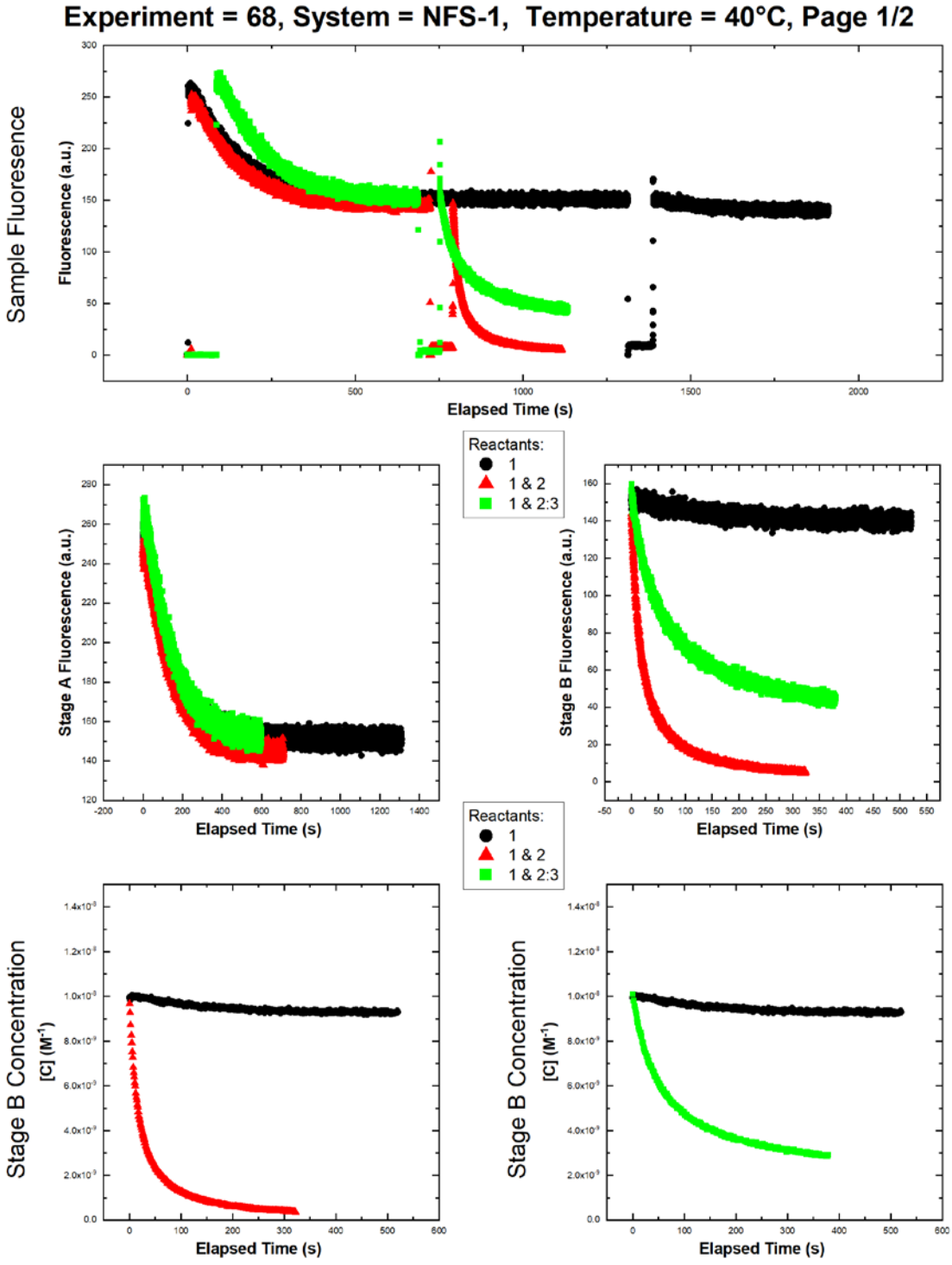

Experiment = 65, System = NFS-1, Temperature = 10°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

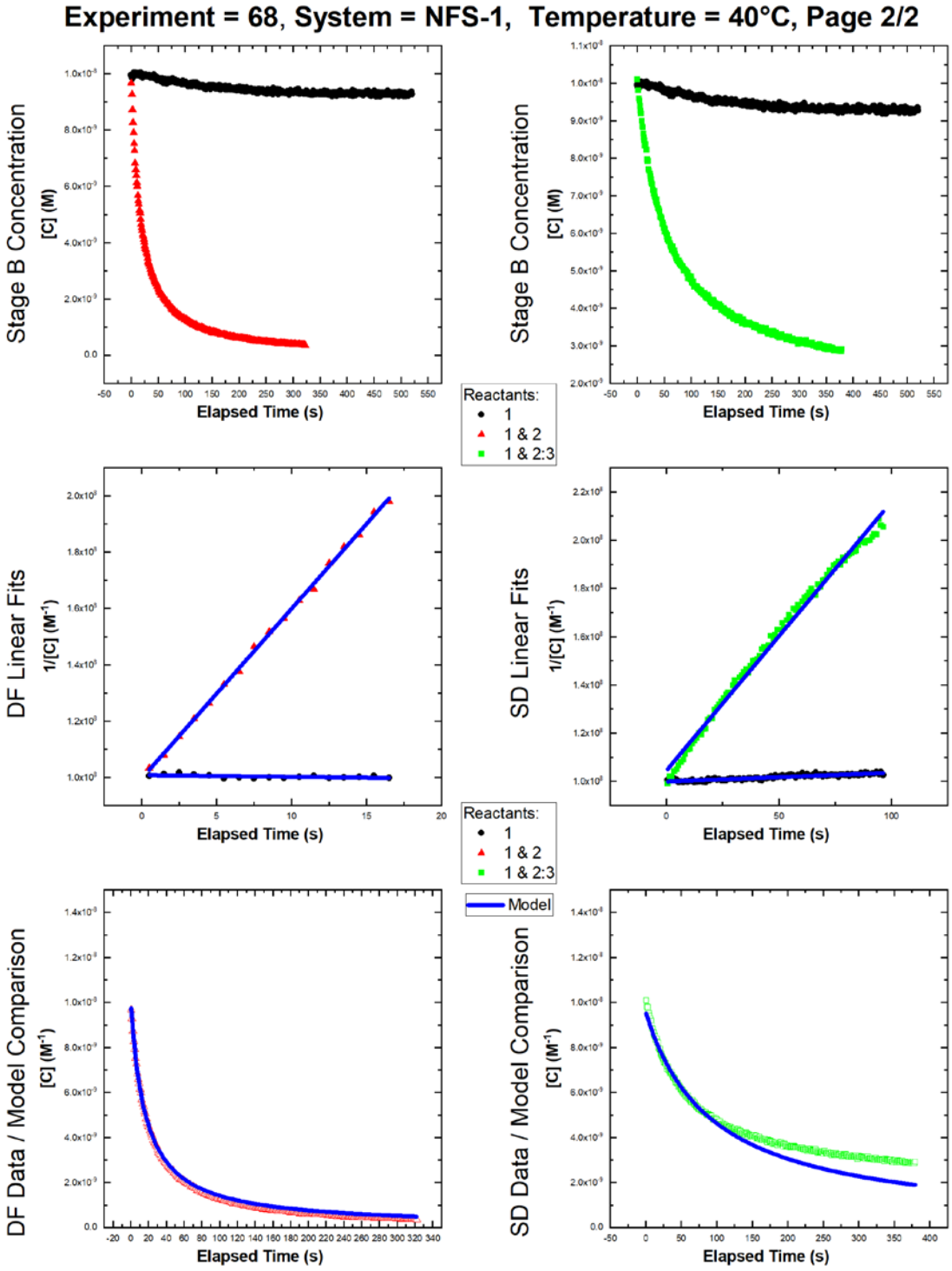

Experiment = 66, System = NFS-1, Temperature = 20°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

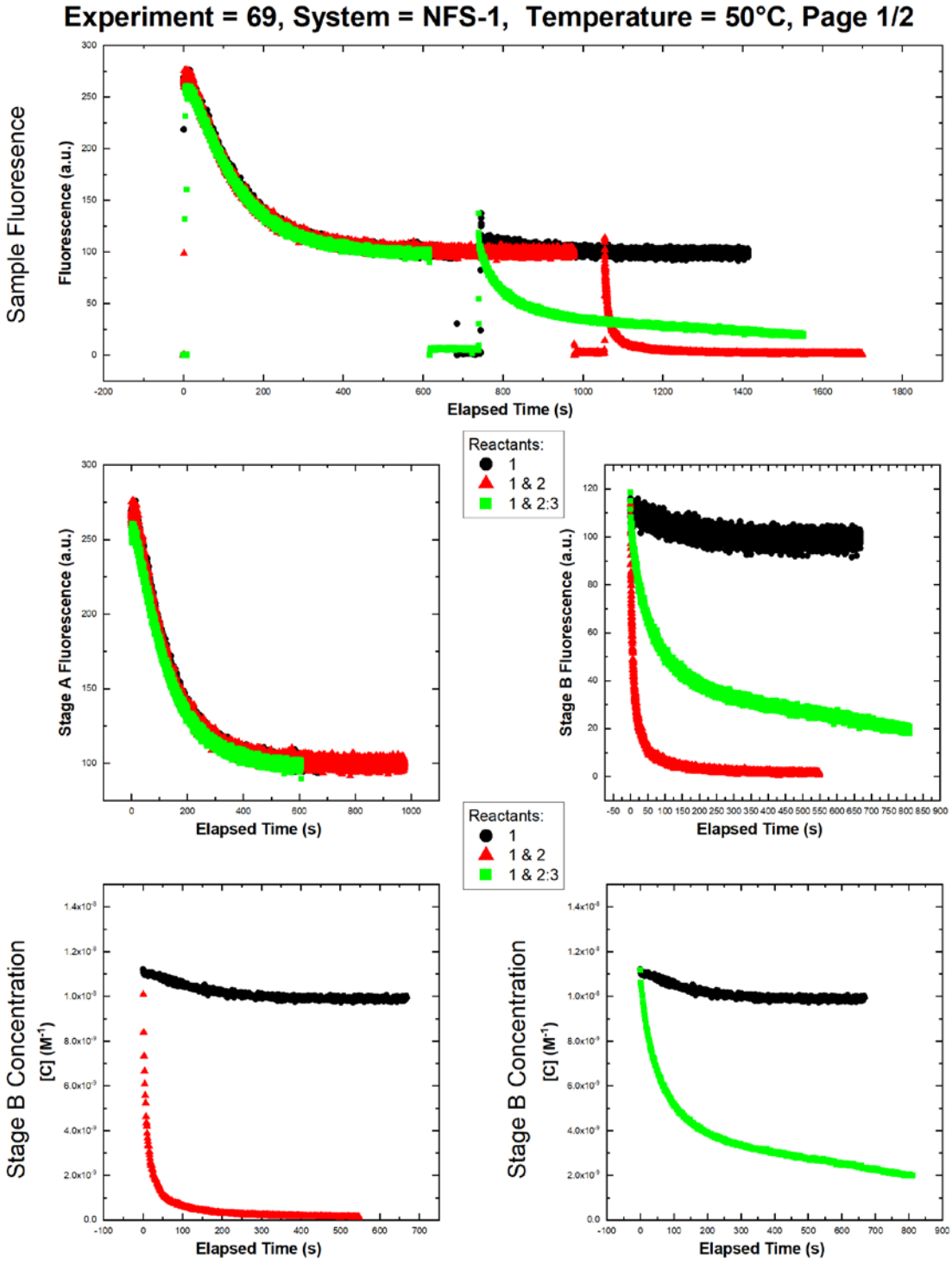

Experiment = 66, System = NFS-1, Temperature = 20°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

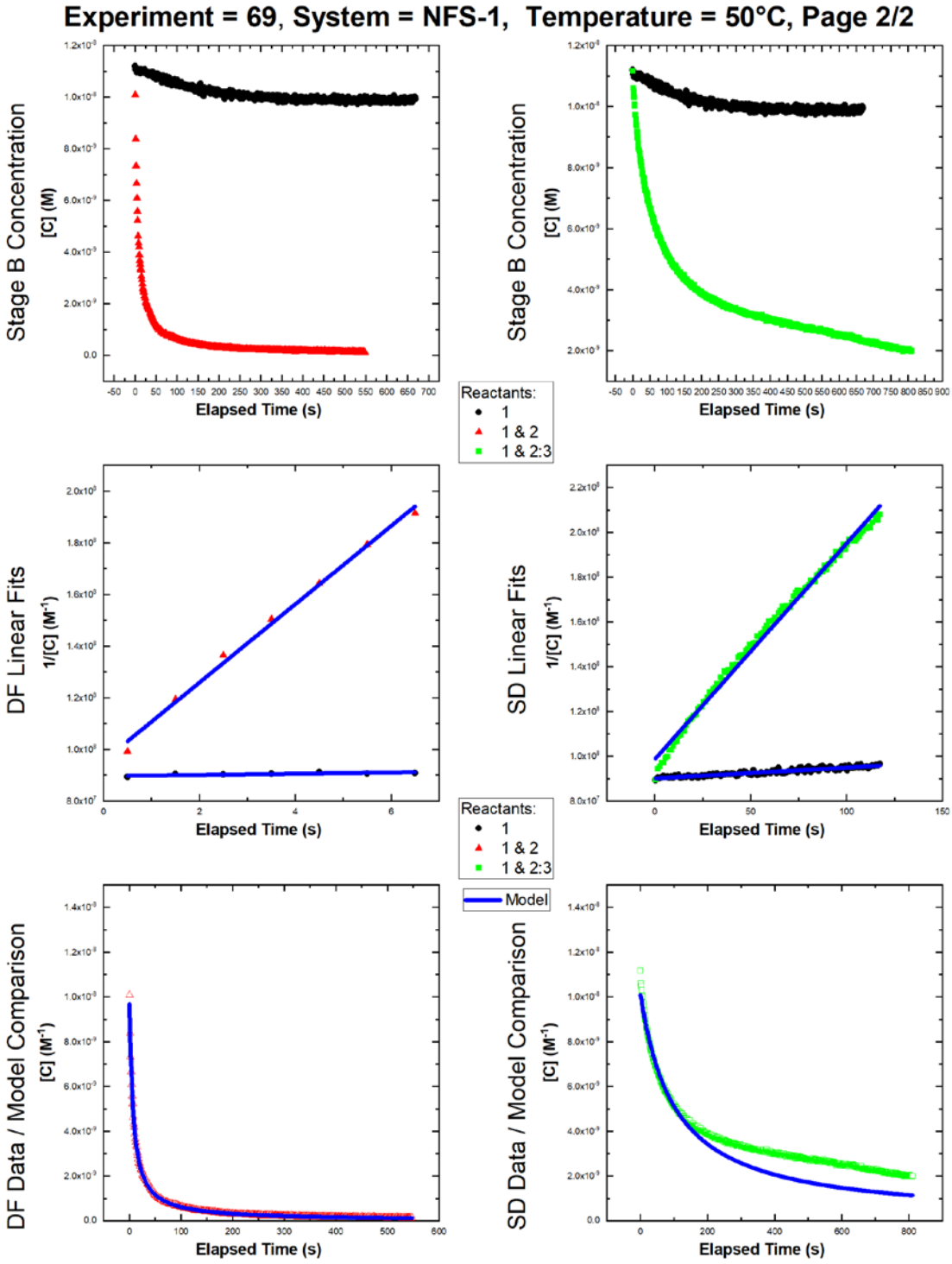

Experiment = 67, System = NFS-1, Temperature = 30°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

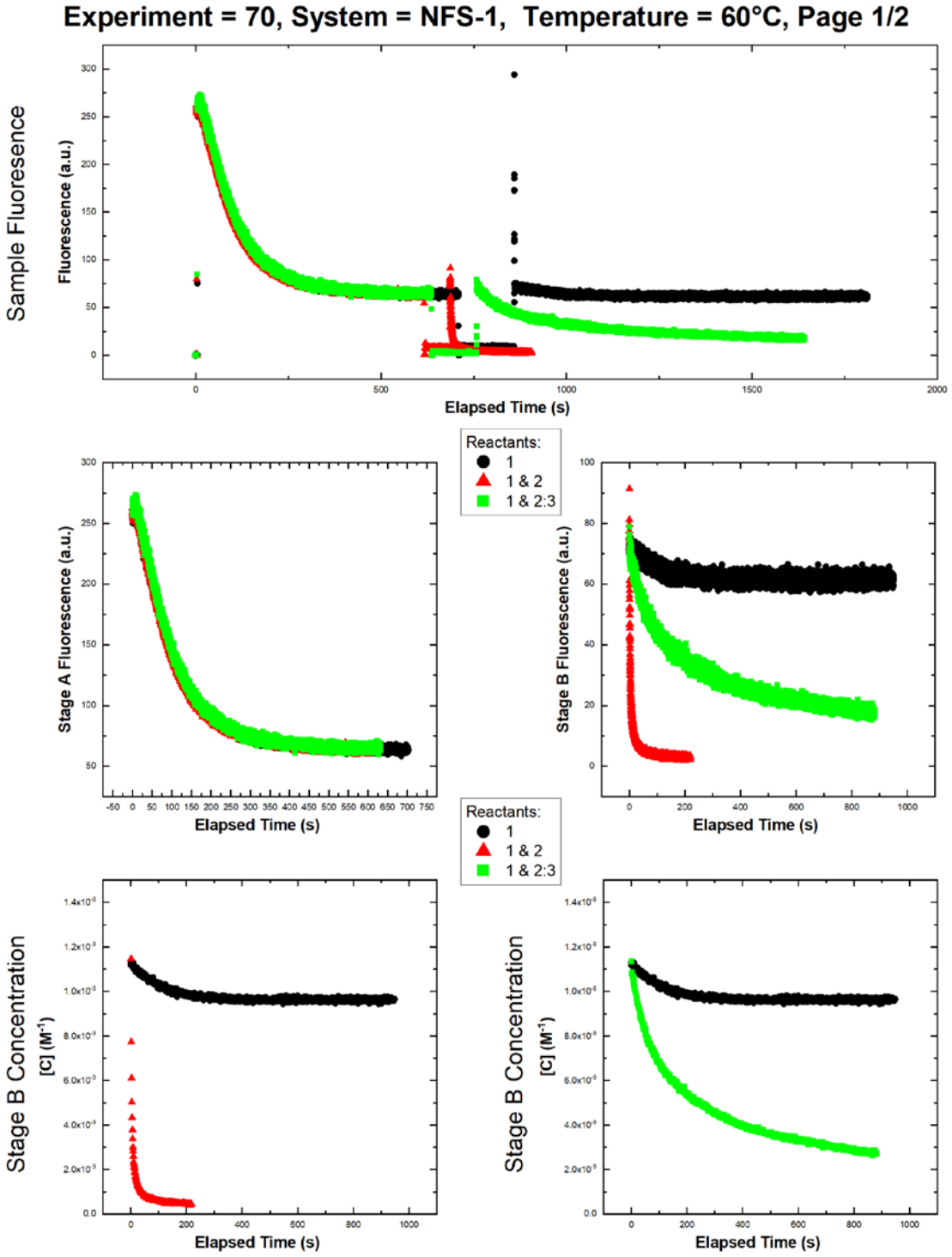

Experiment = 67, System = NFS-1, Temperature = 30°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

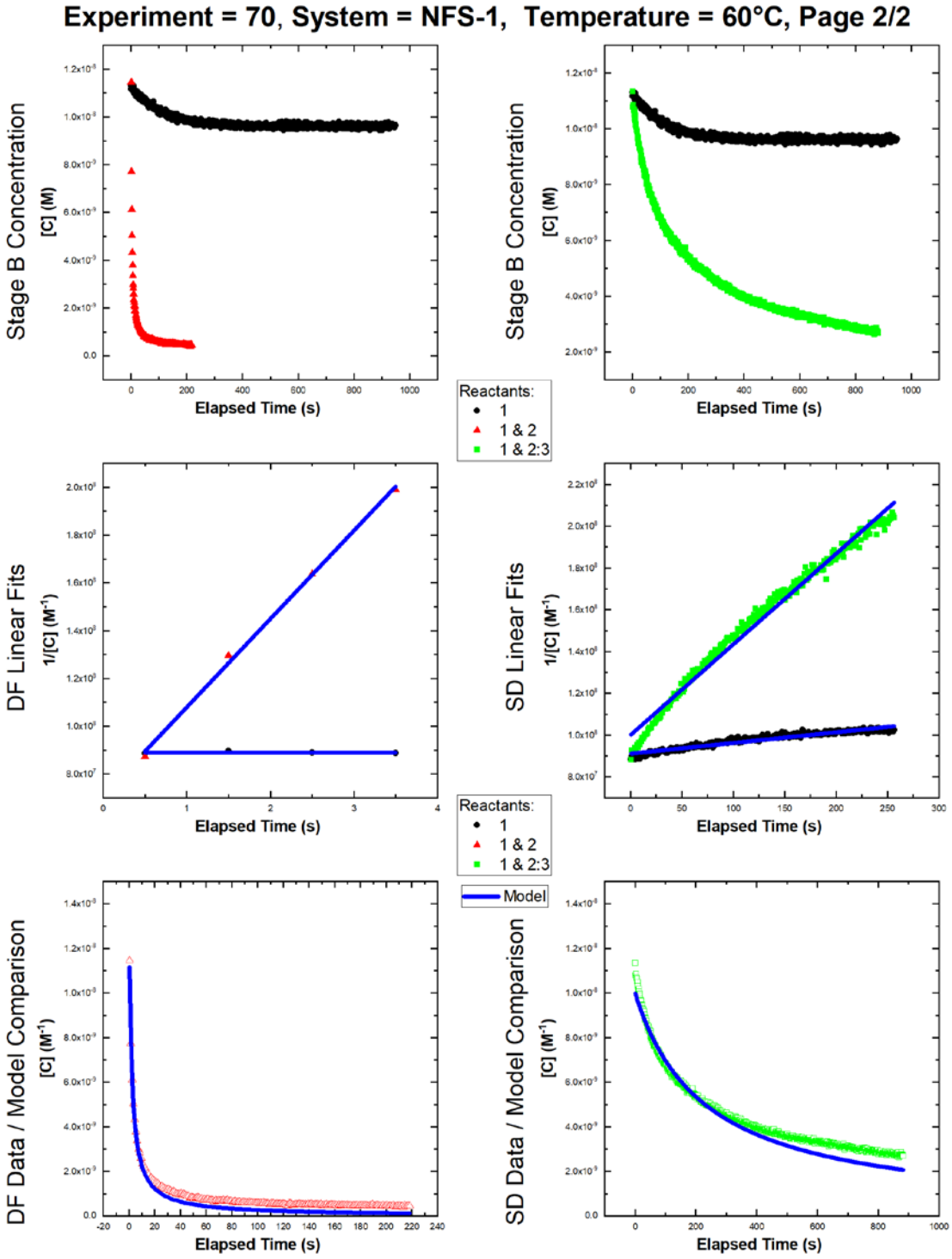

Experiment = 68, System = NFS-1, Temperature = 40°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

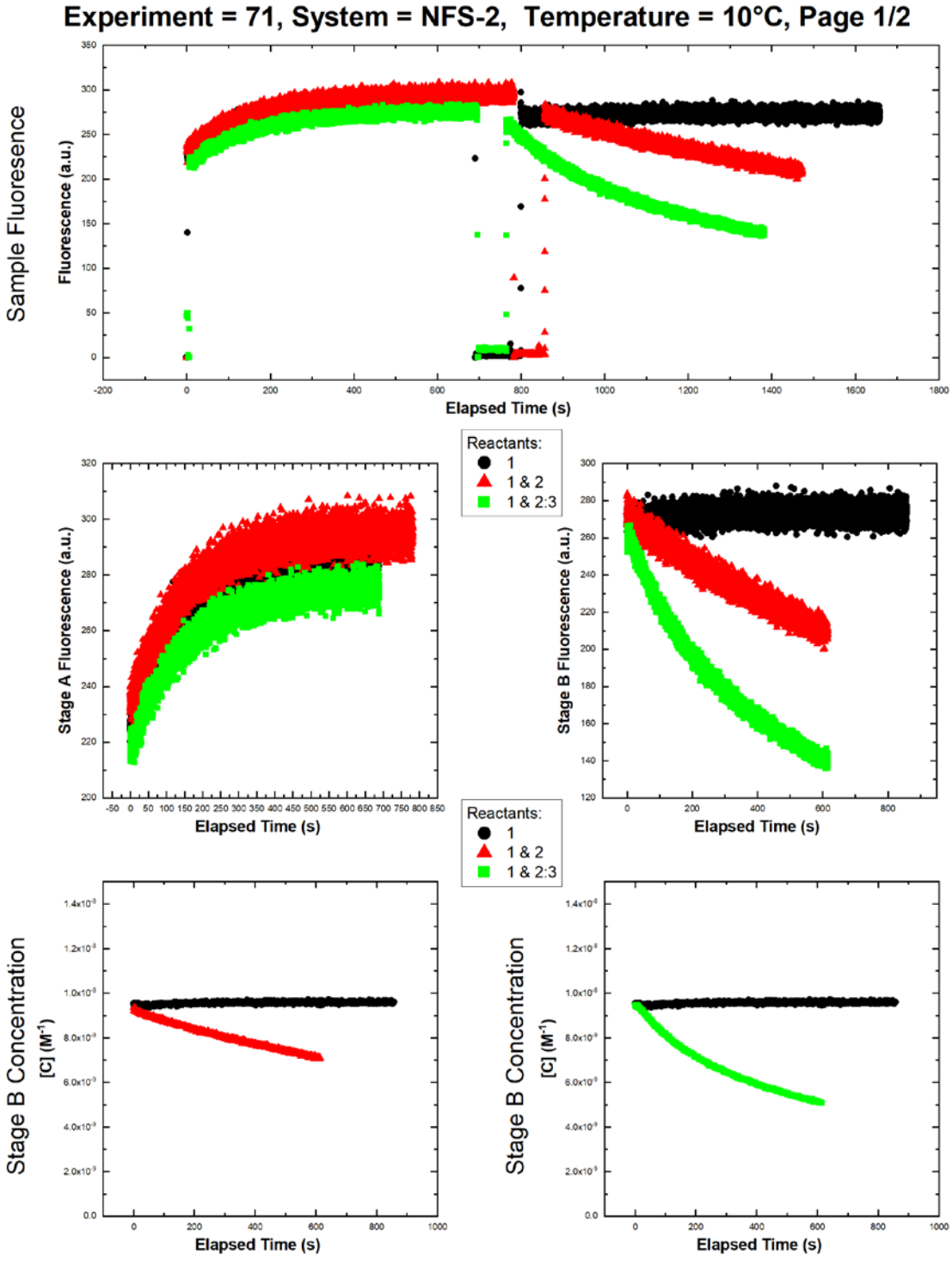

Experiment = 68, System = NFS-1, Temperature = 40°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

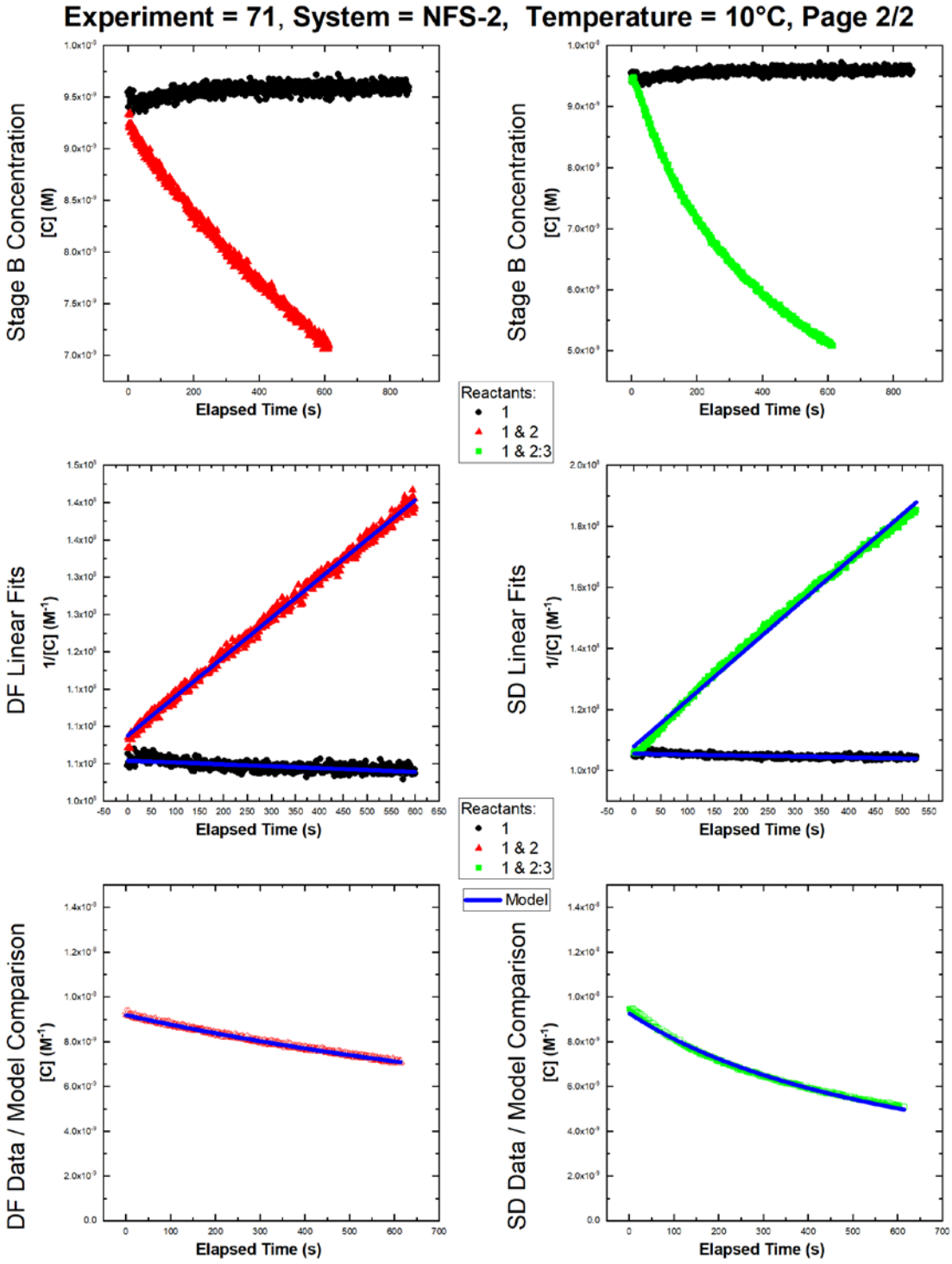

Experiment = 69, System = NFS-1, Temperature = 50°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

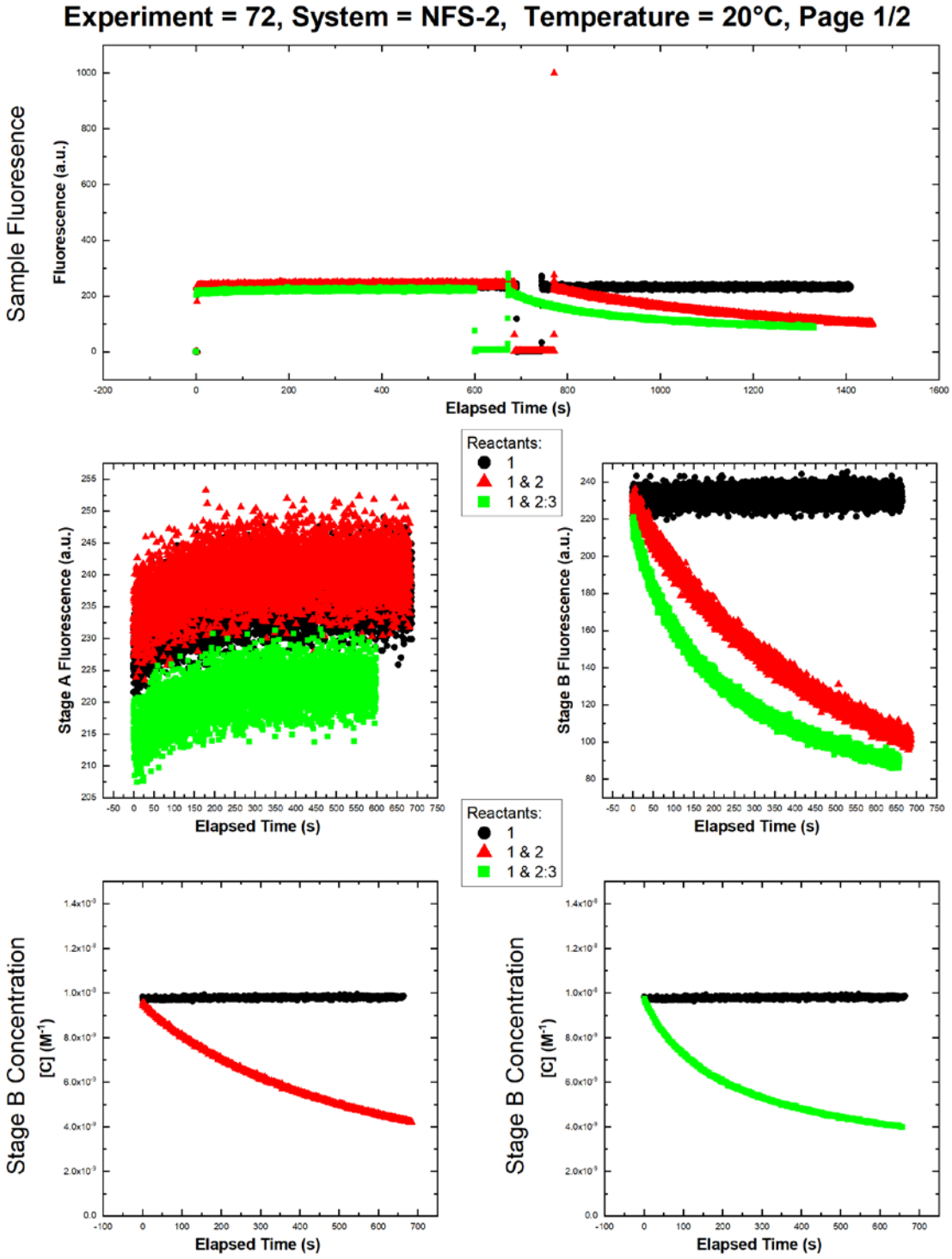

Experiment = 69, System = NFS-1, Temperature = 50°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

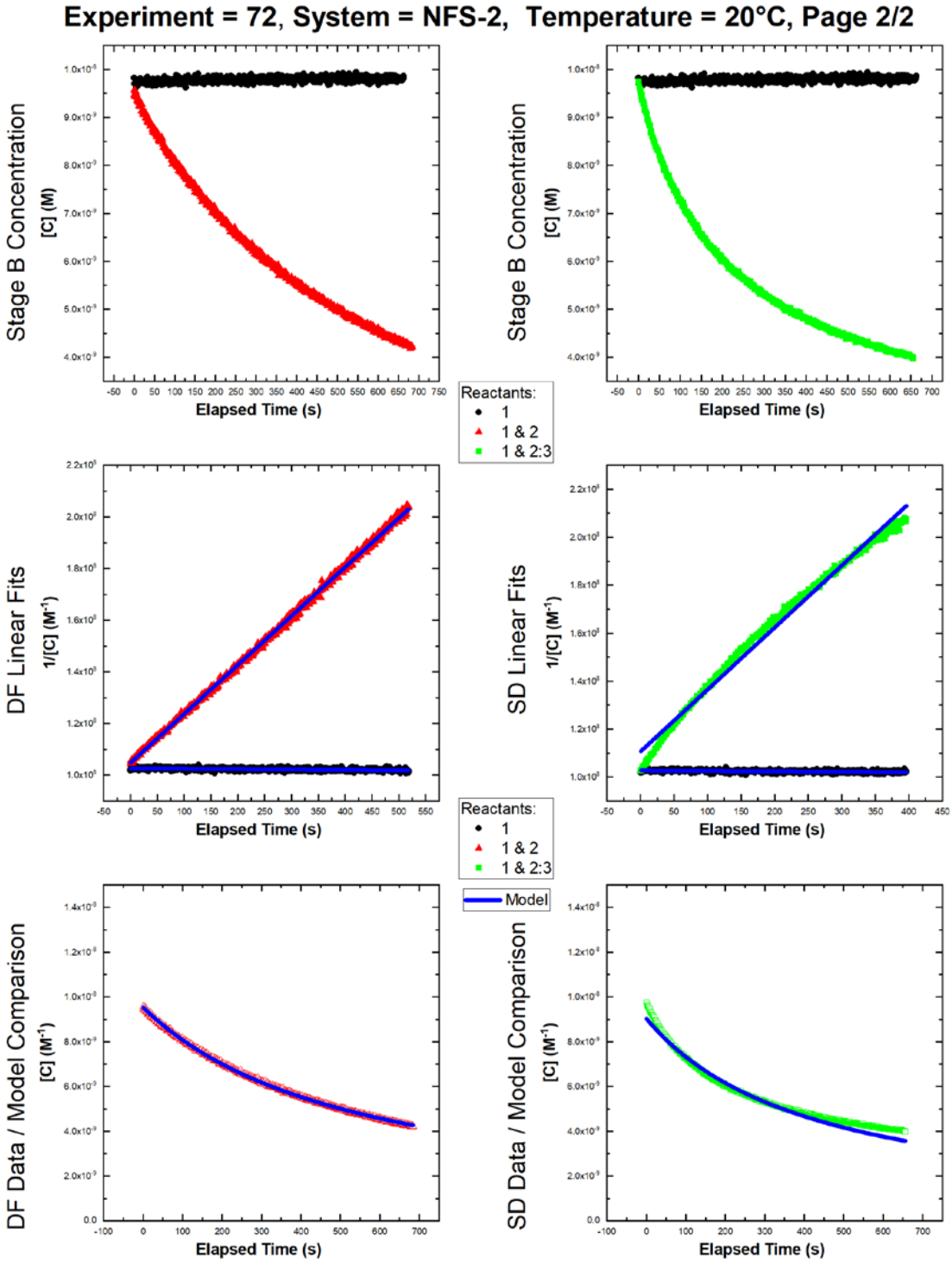

Experiment = 70, System = NFS-1, Temperature = 60°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

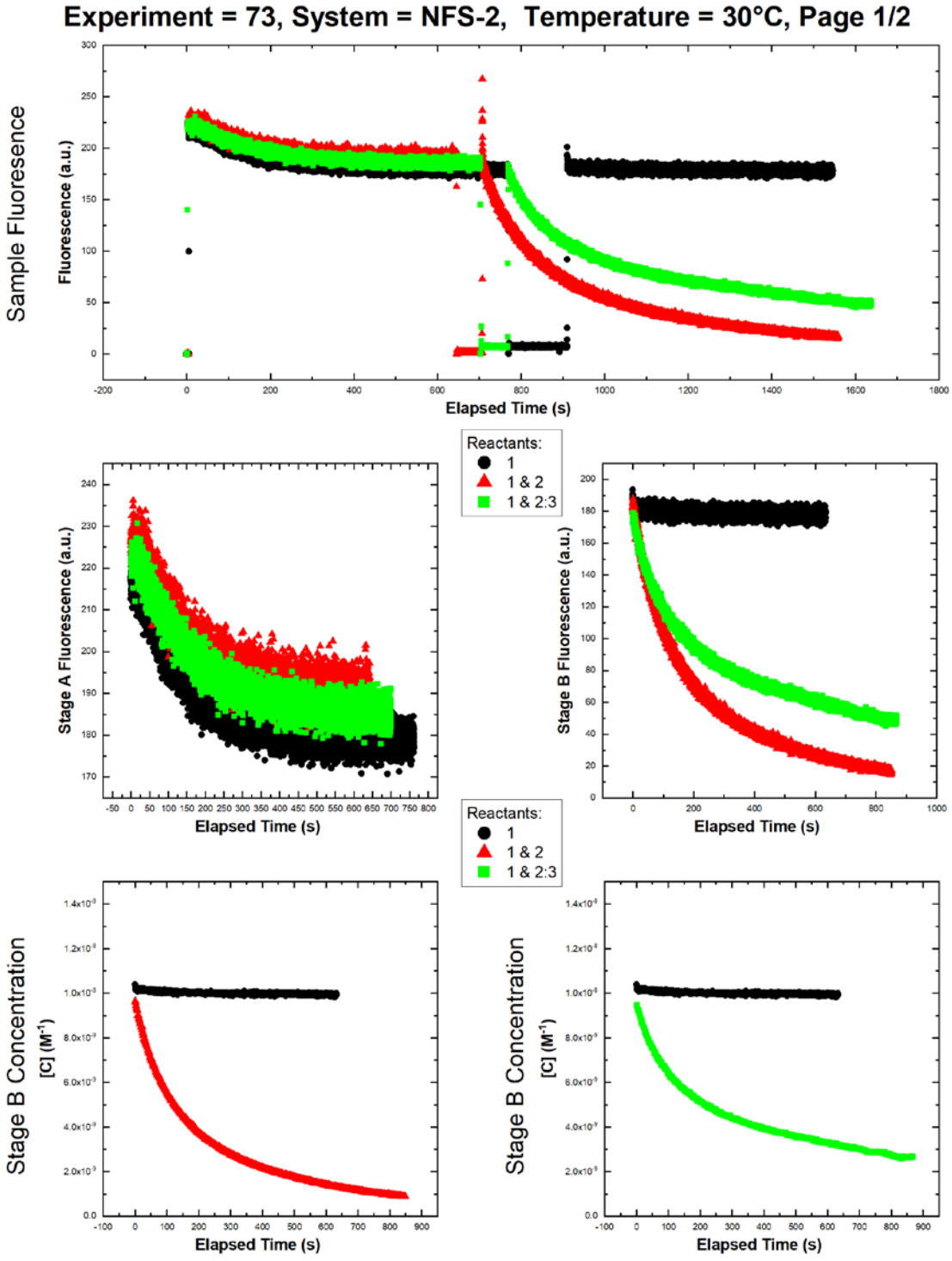

Experiment = 70, System = NFS-1, Temperature = 60°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
[C] (M-1)
SD Data / Model Comparison

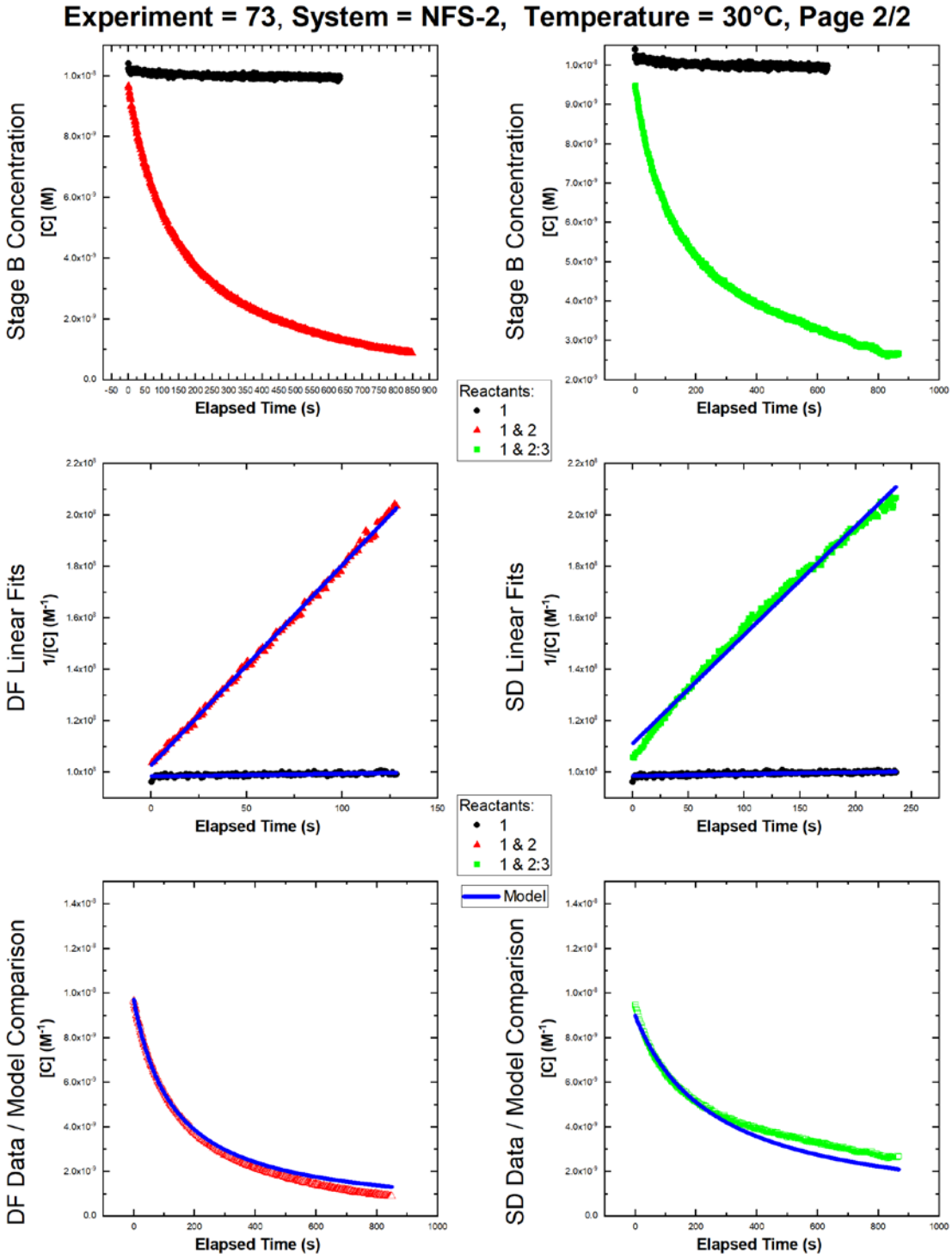

Experiment = 71, System = NFS-2, Temperature = 10°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

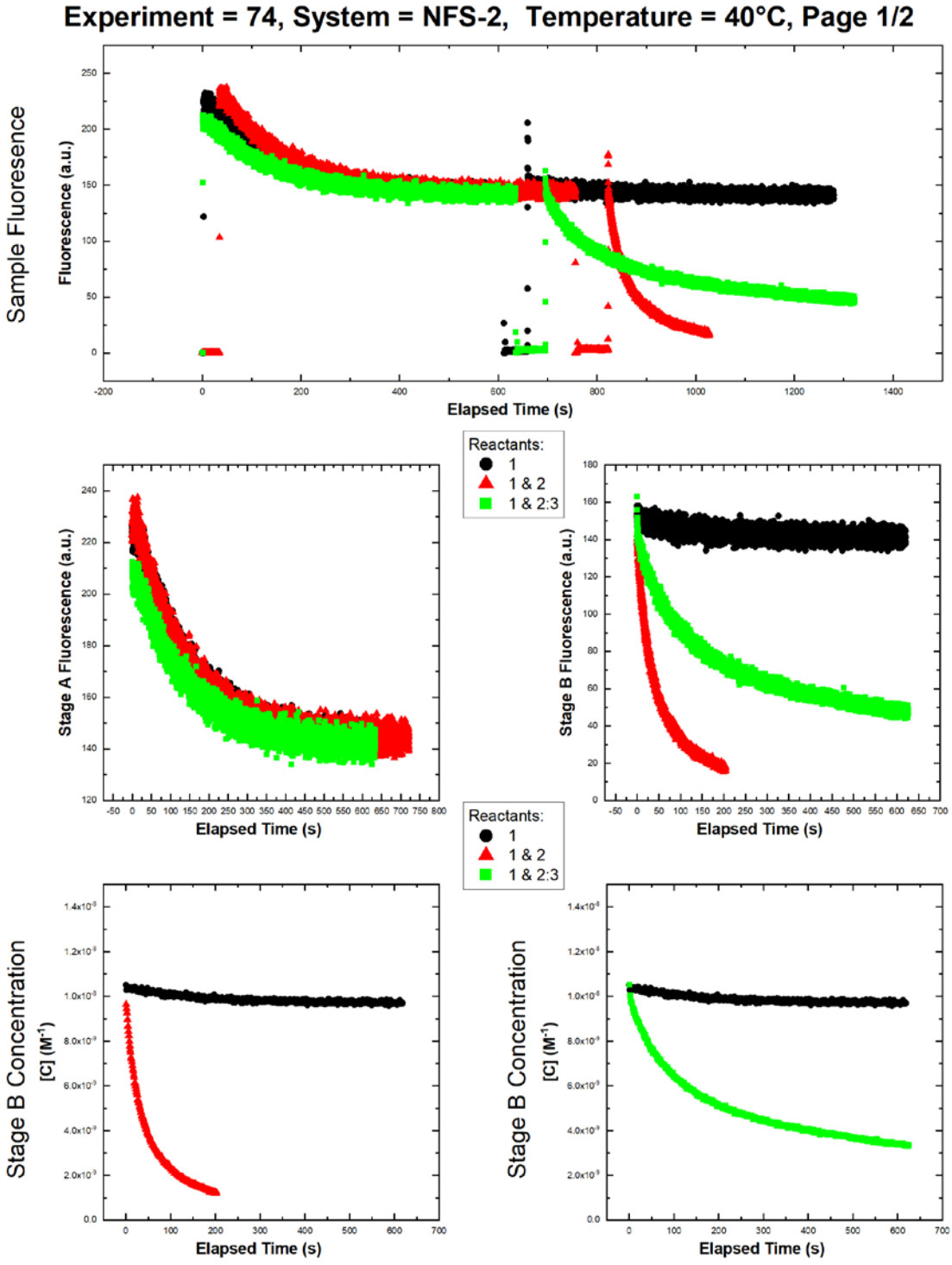

Experiment = 71, System = NFS-2, Temperature = 10°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

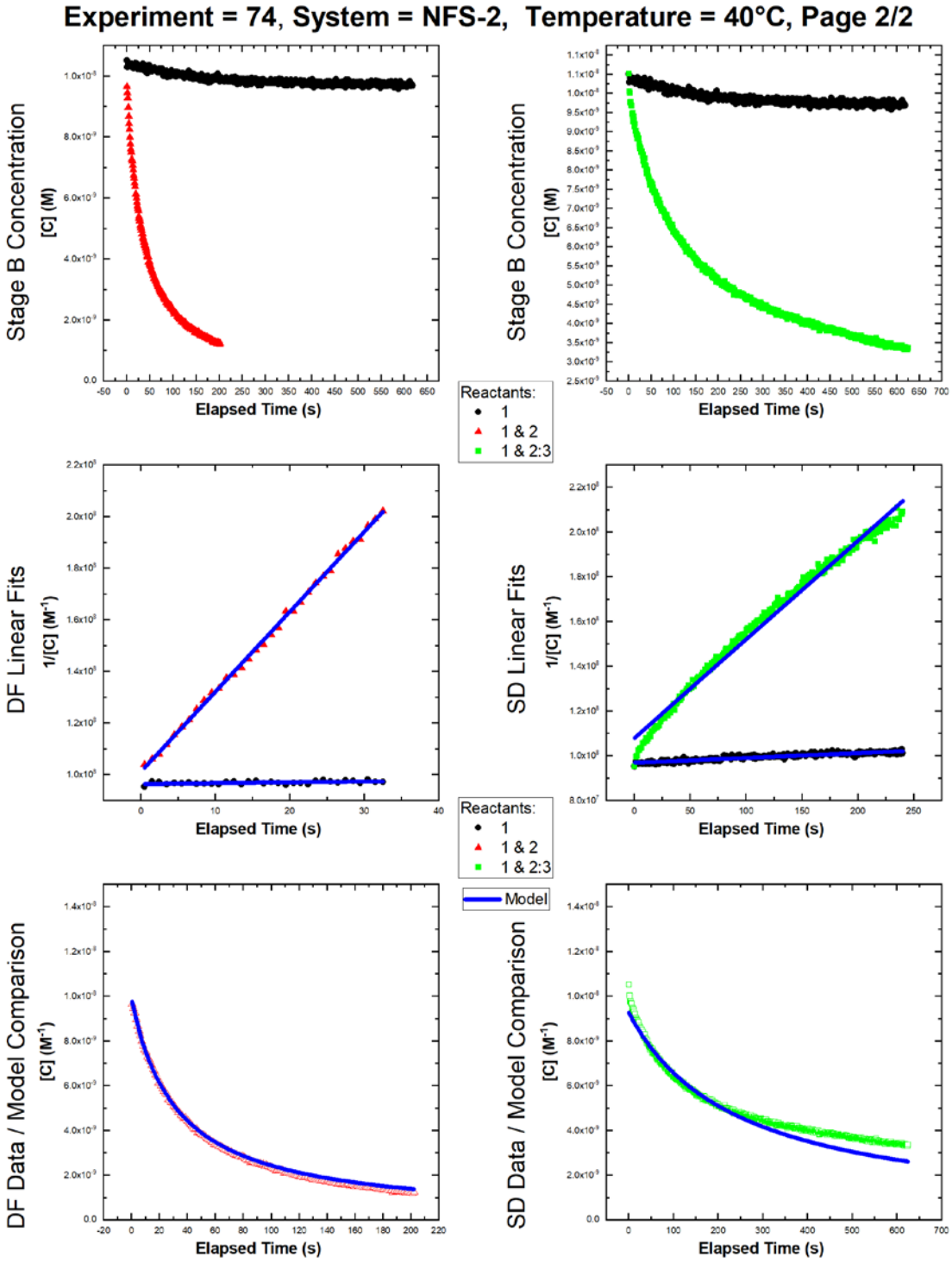

Experiment = 72, System = NFS-2, Temperature = 20°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

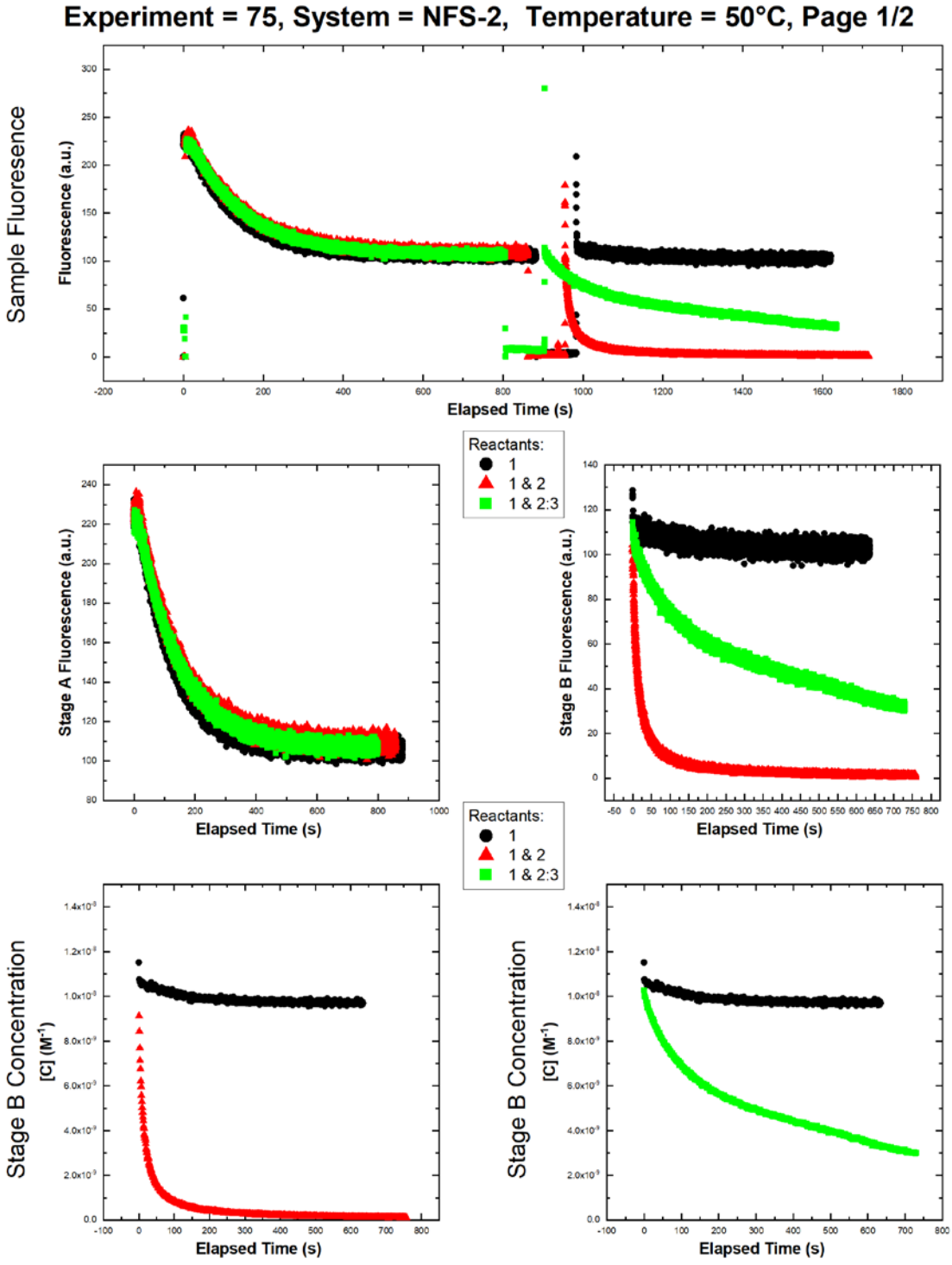

Experiment = 72, System = NFS-2, Temperature = 20°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

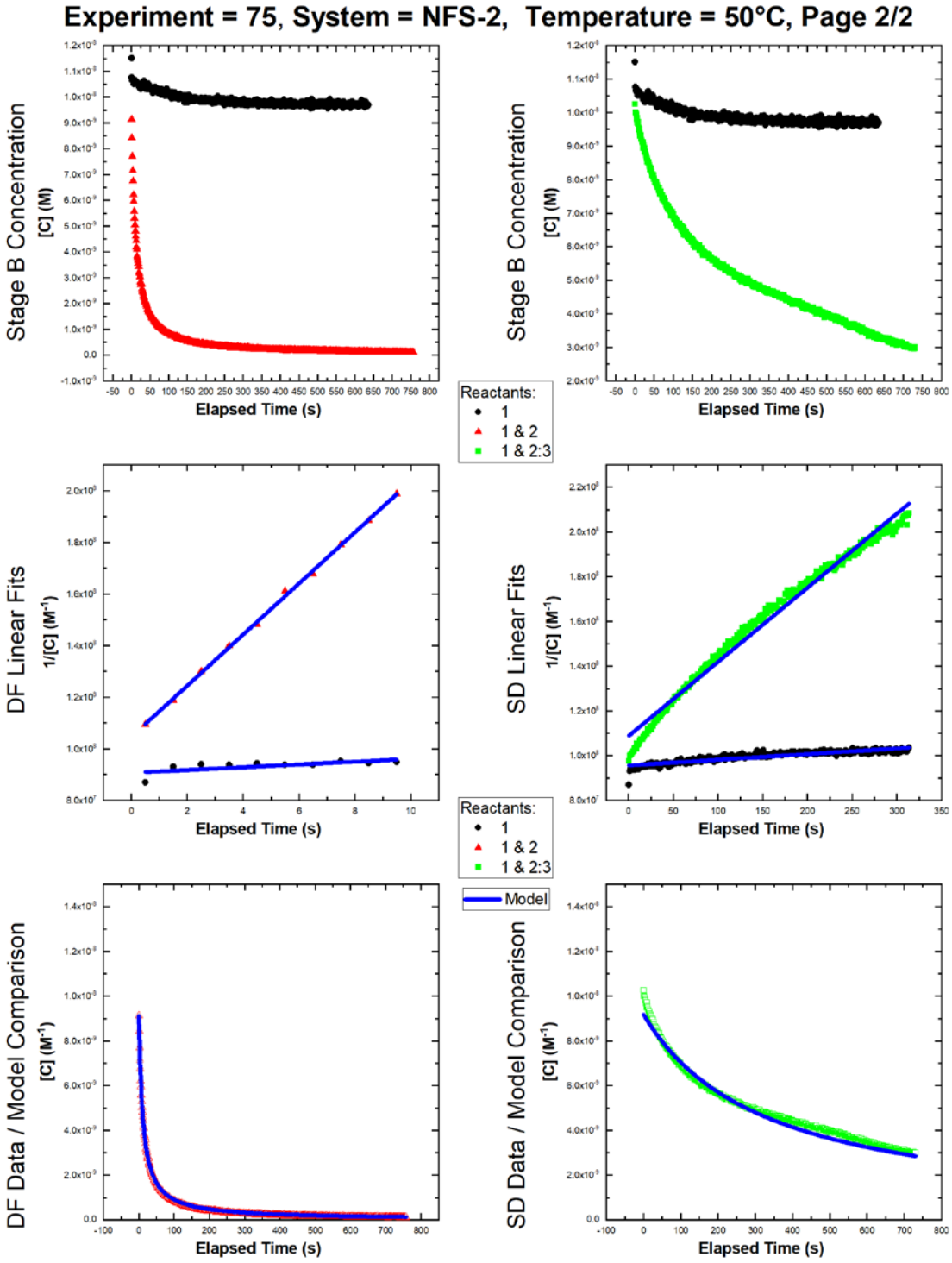

Experiment = 73, System = NFS-2, Temperature = 30°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

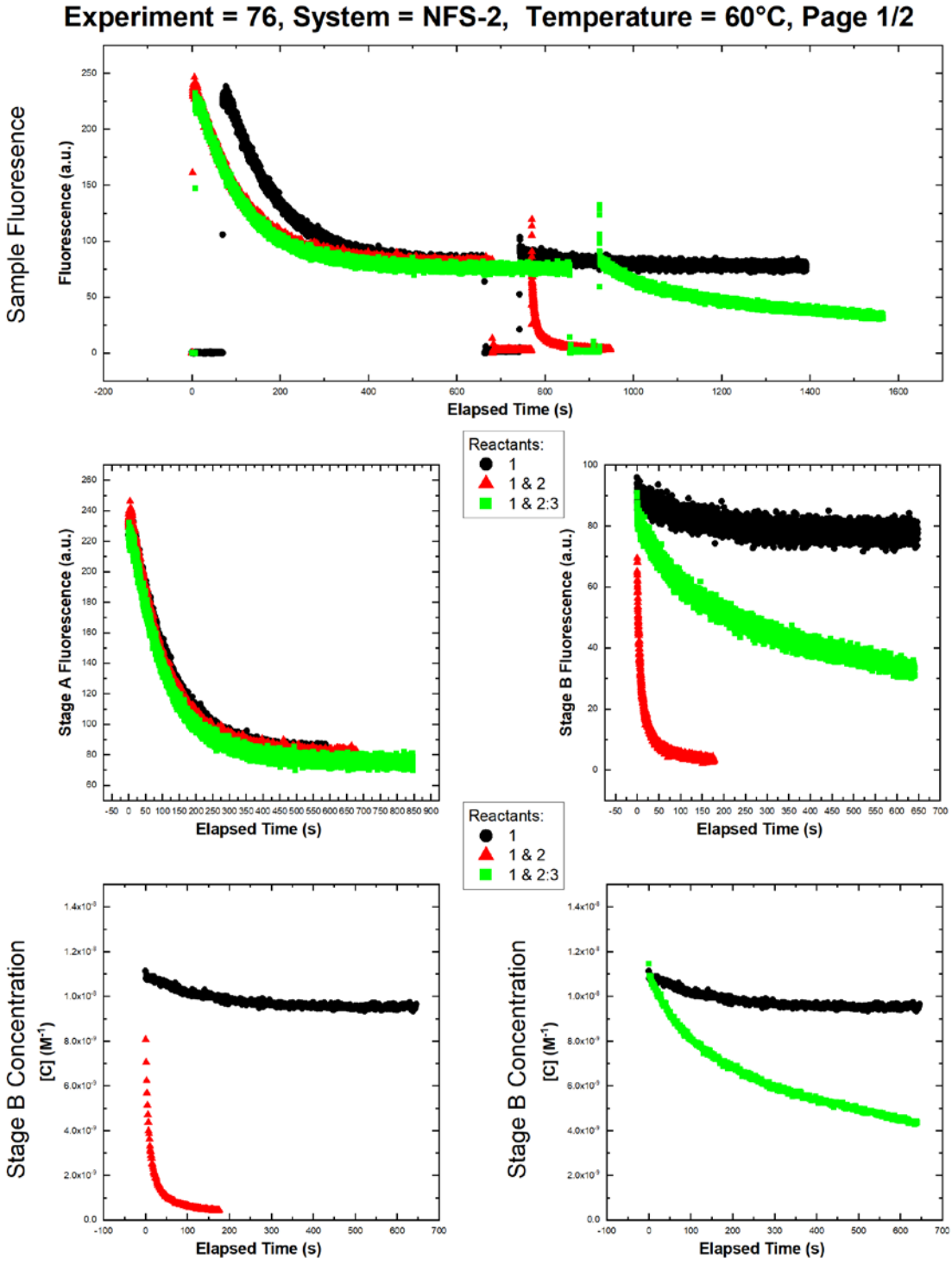

Experiment = 73, System = NFS-2, Temperature = 30°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

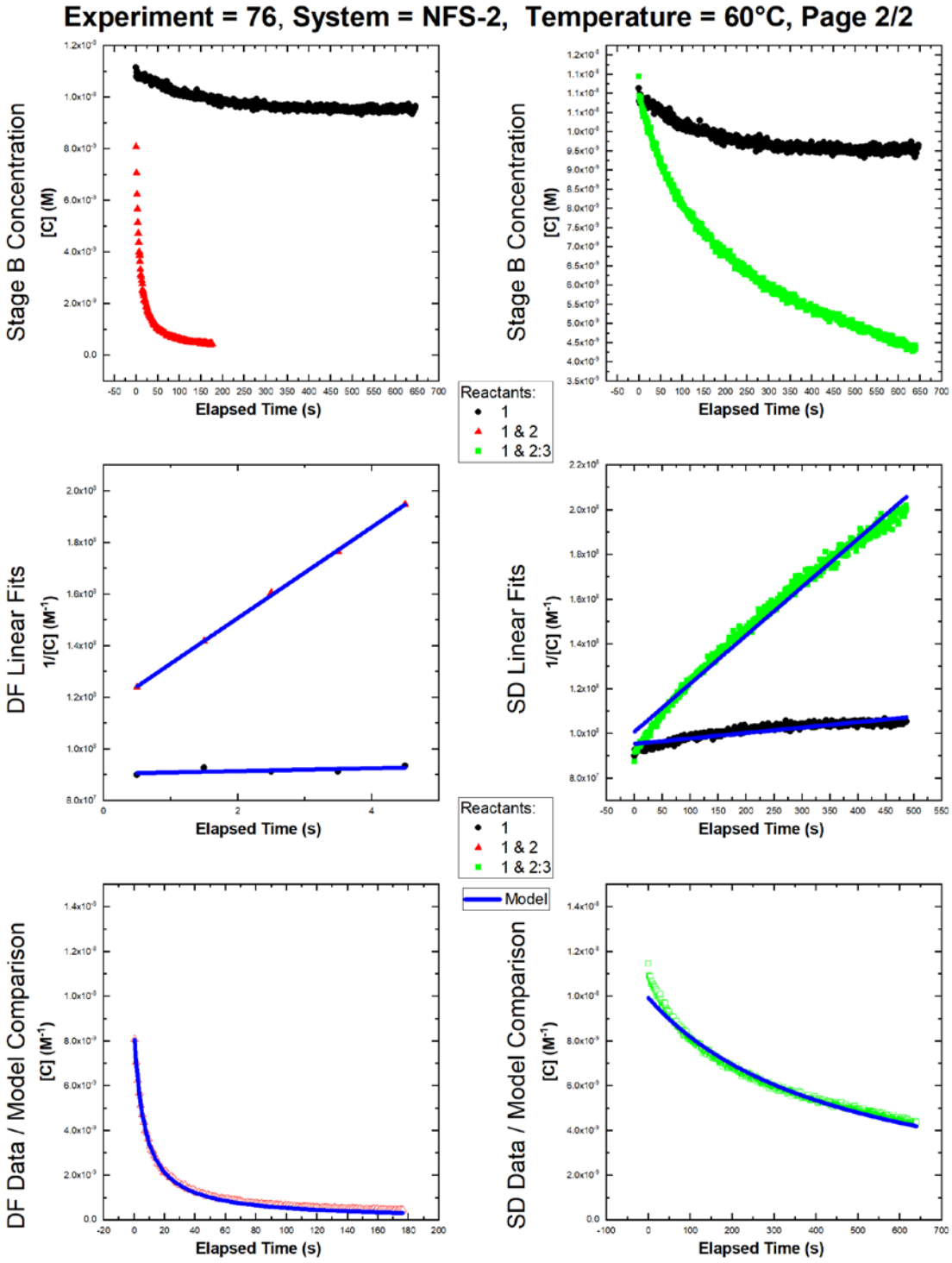

Experiment = 74, System = NFS-2, Temperature = 40°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

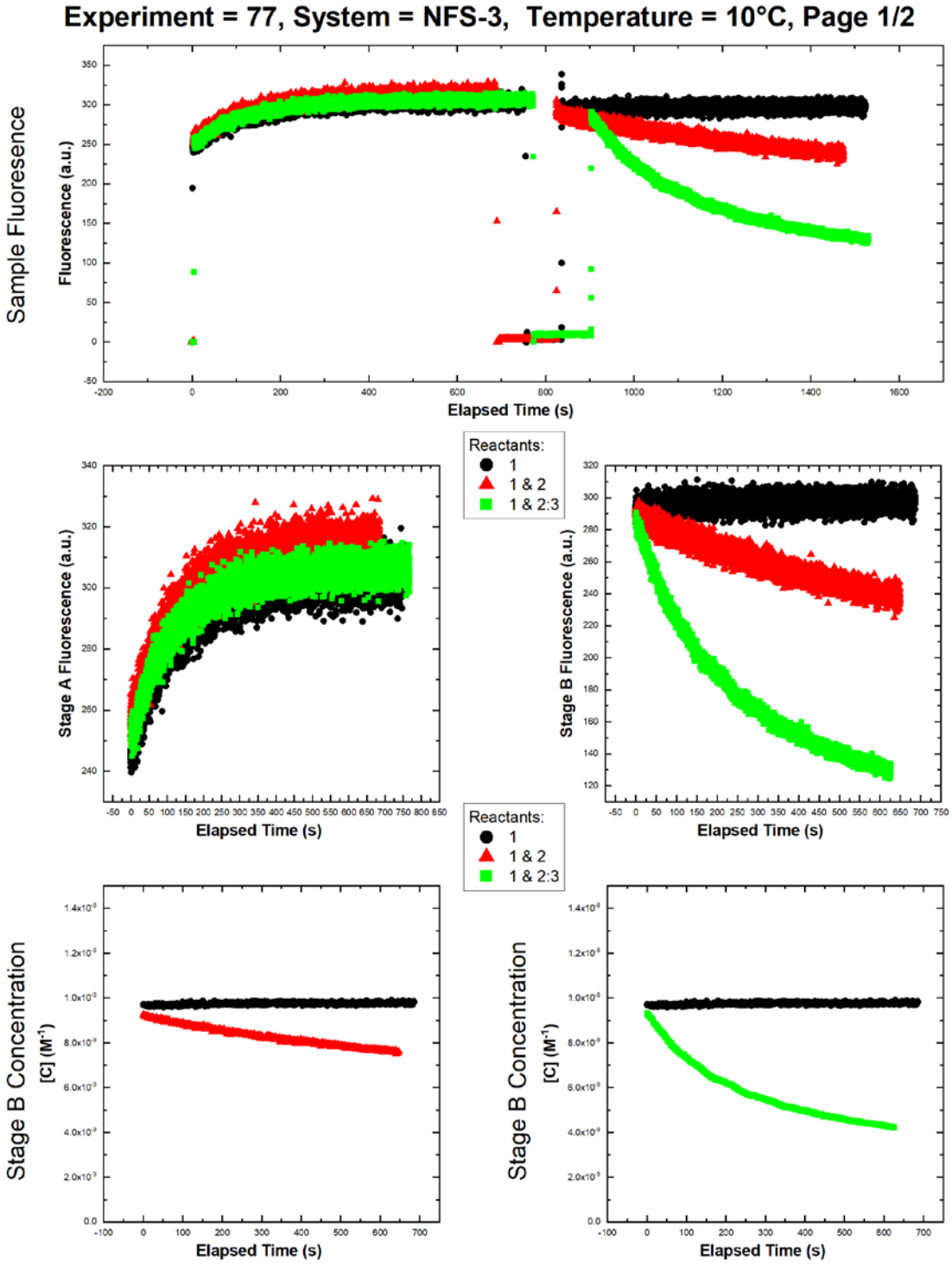

Experiment = 74, System = NFS-2, Temperature = 40°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

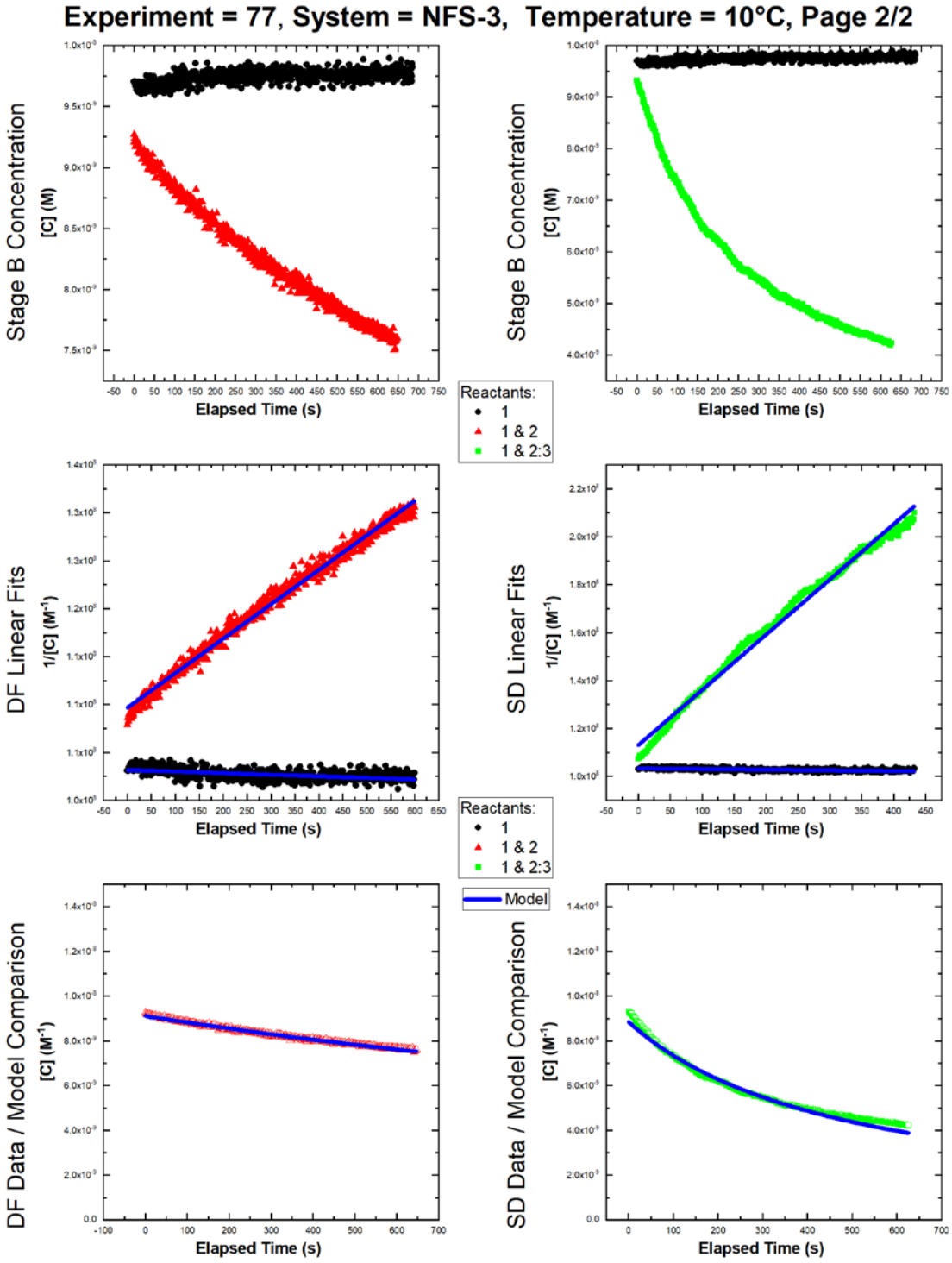

Experiment = 75, System = NFS-2, Temperature = 50°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

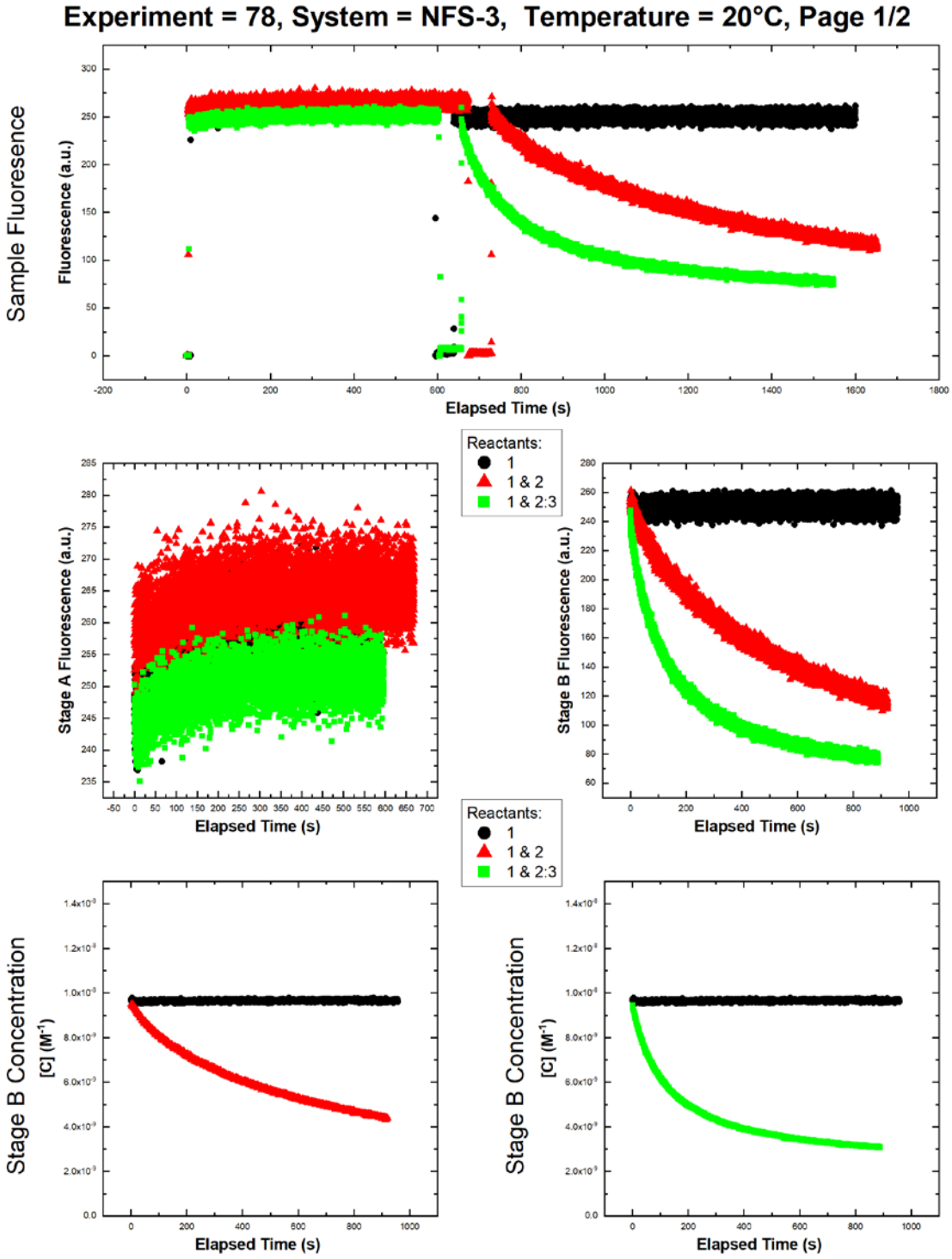

Experiment = 75, System = NFS-2, Temperature = 50°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

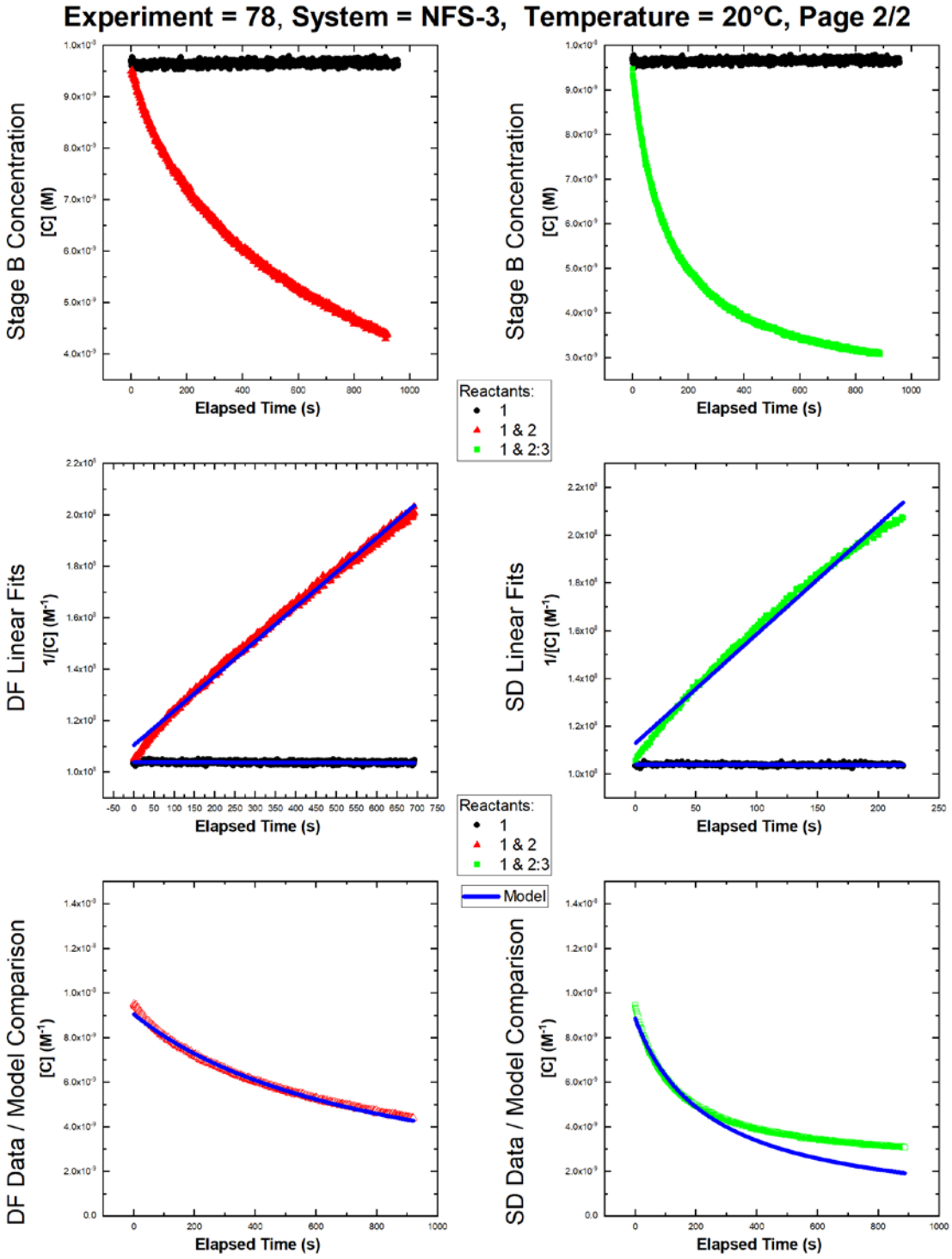

Experiment = 76, System = NFS-2, Temperature = 60°C, Page 1/2
Sample Fluorescence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

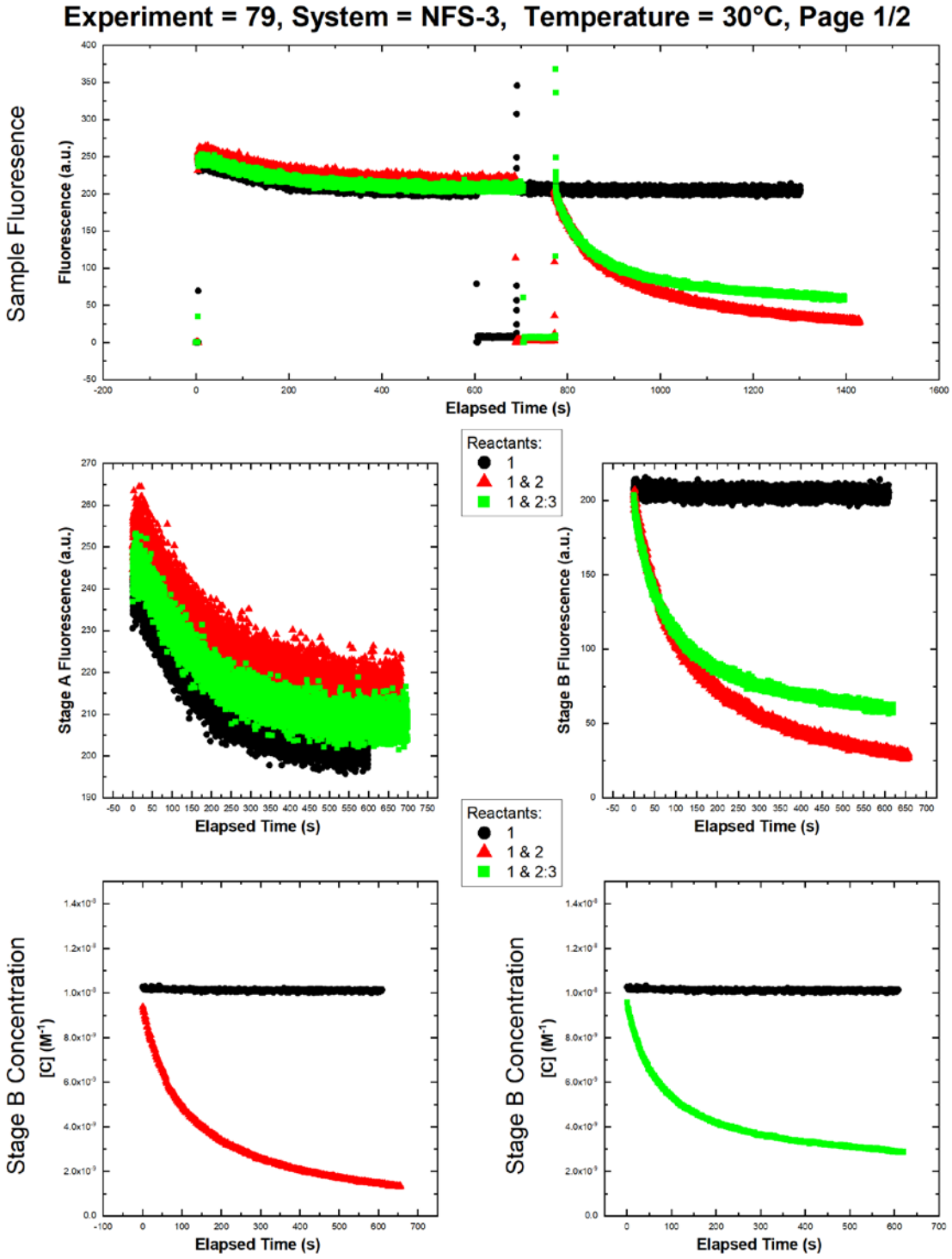

Experiment = 76, System = NFS-2, Temperature = 60°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
[C] (M-1)
SD Data / Model Comparison

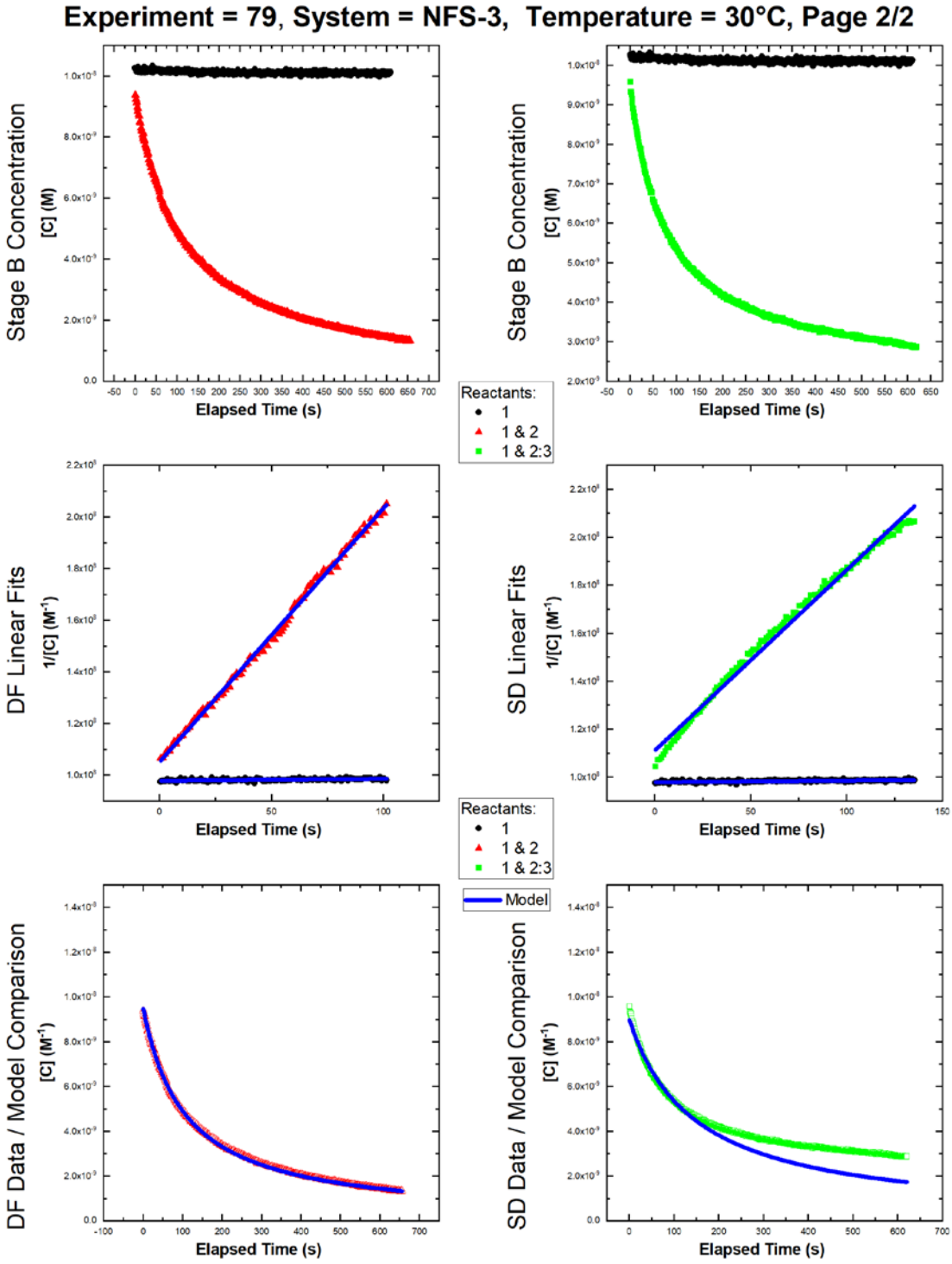

Experiment = 77, System = NFS-3, Temperature = 10°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

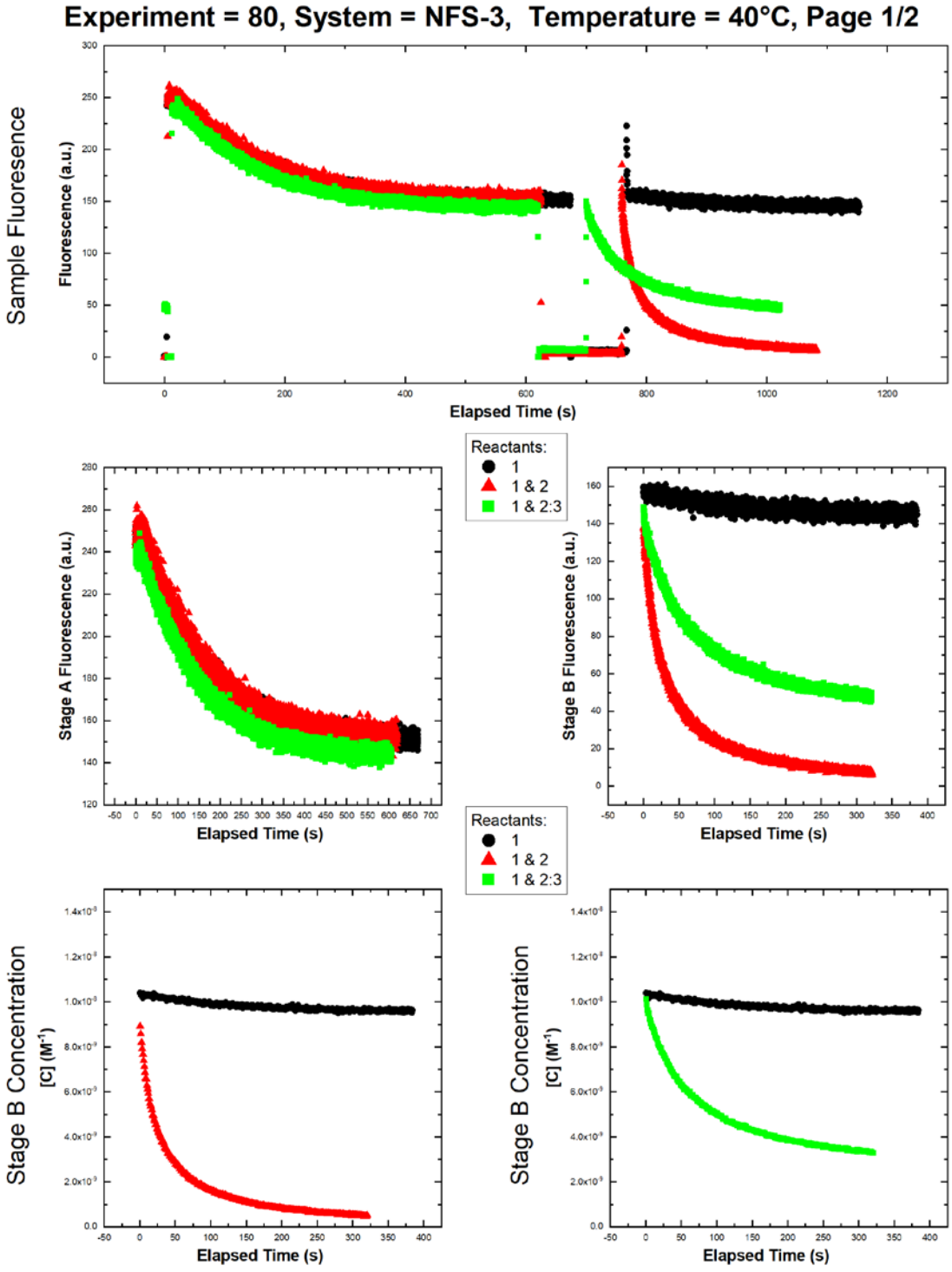

Experiment = 77, System = NFS-3, Temperature = 10°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

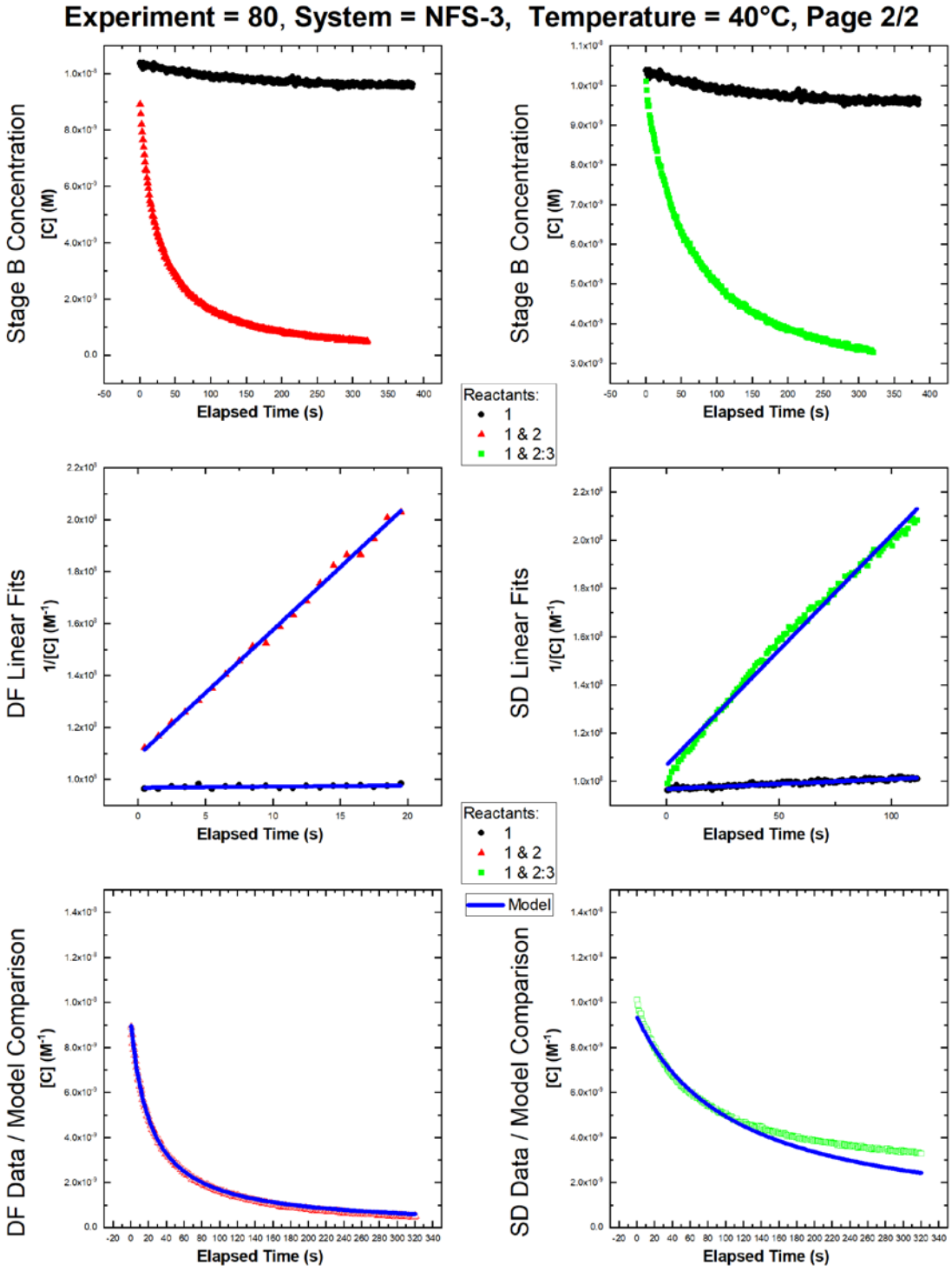

Experiment = 78, System = NFS-3, Temperature = 20°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

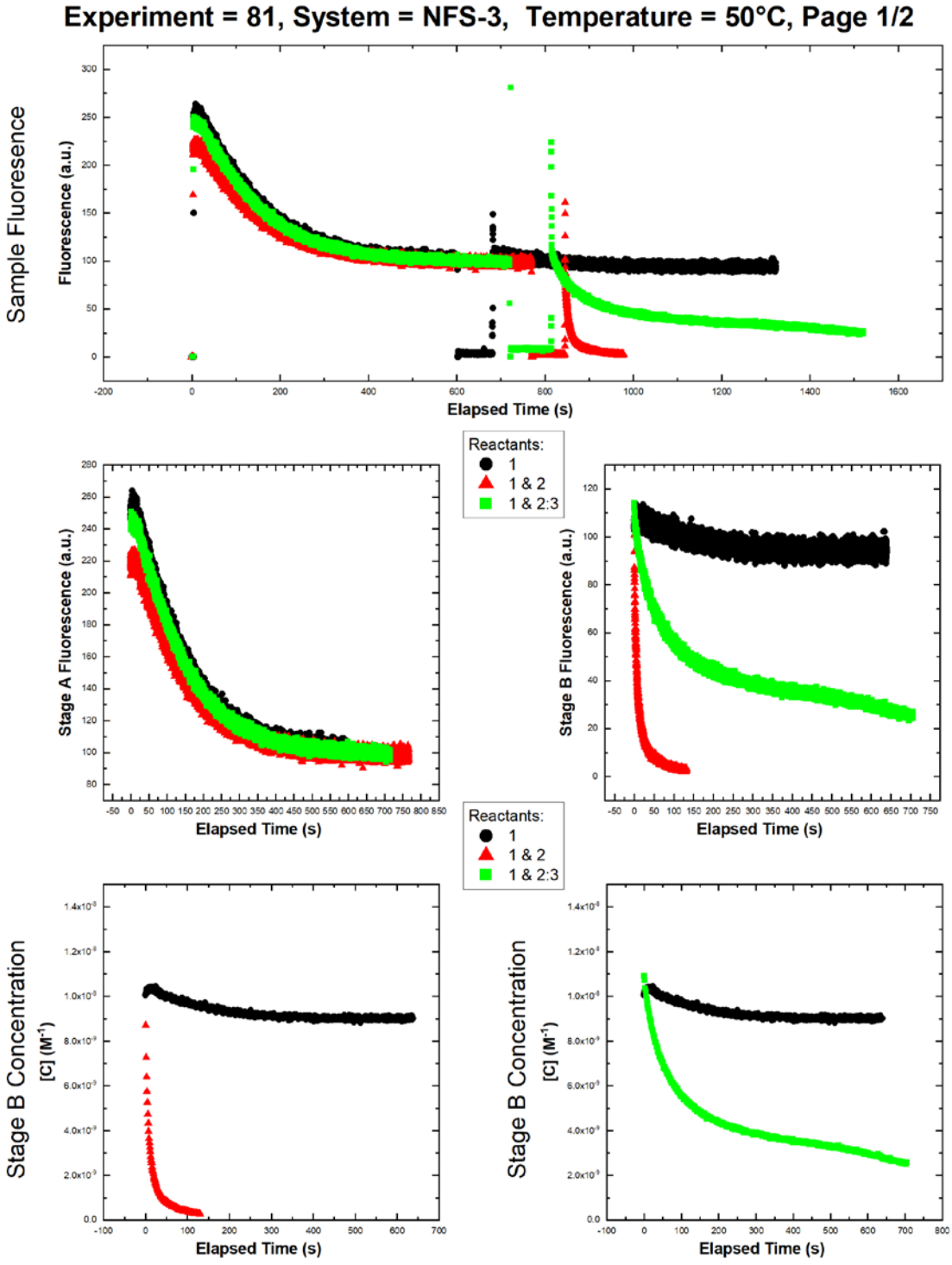

Experiment = 78, System = NFS-3, Temperature = 20°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

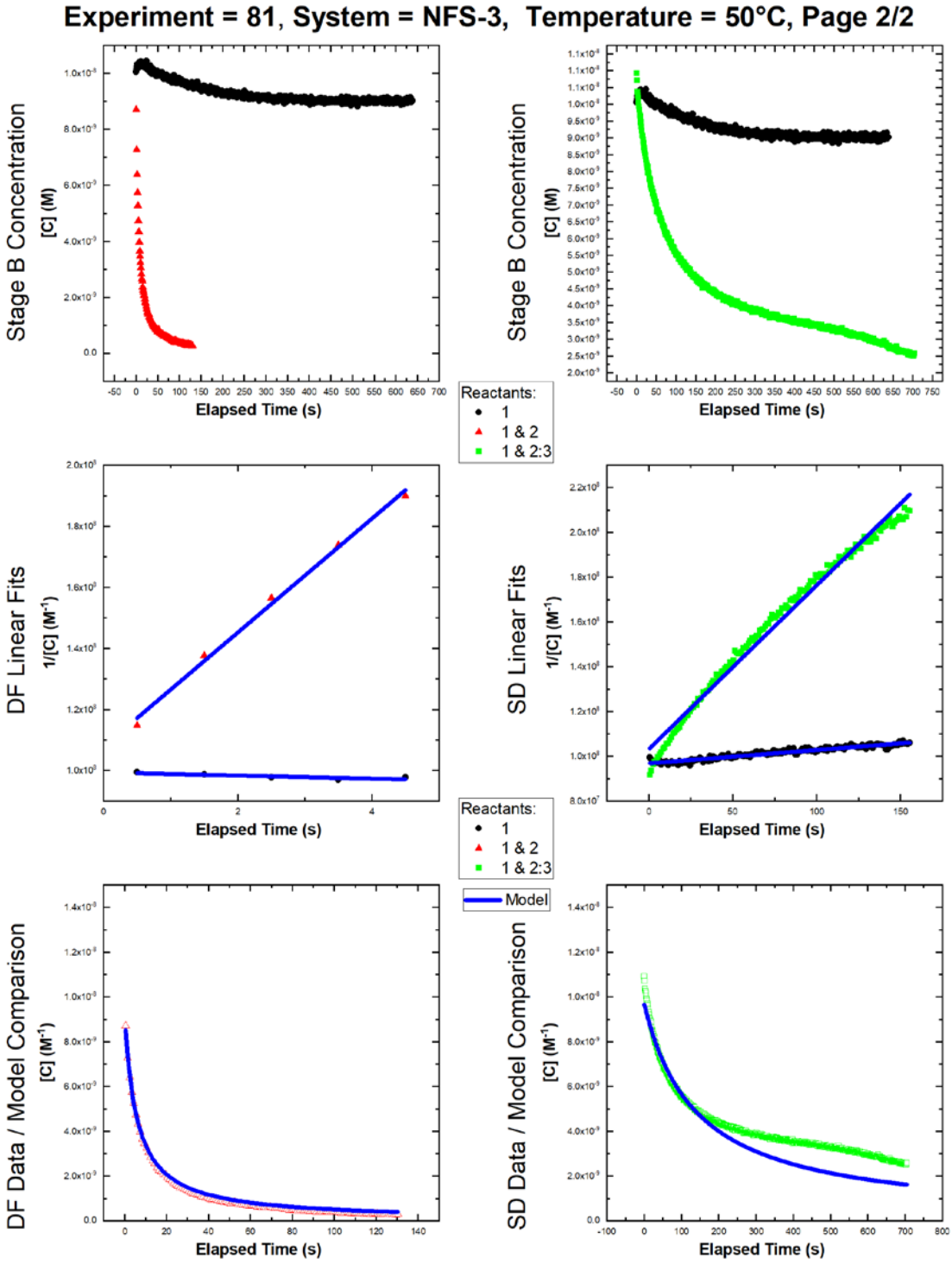

Experiment = 79, System = NFS-3, Temperature = 30°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)

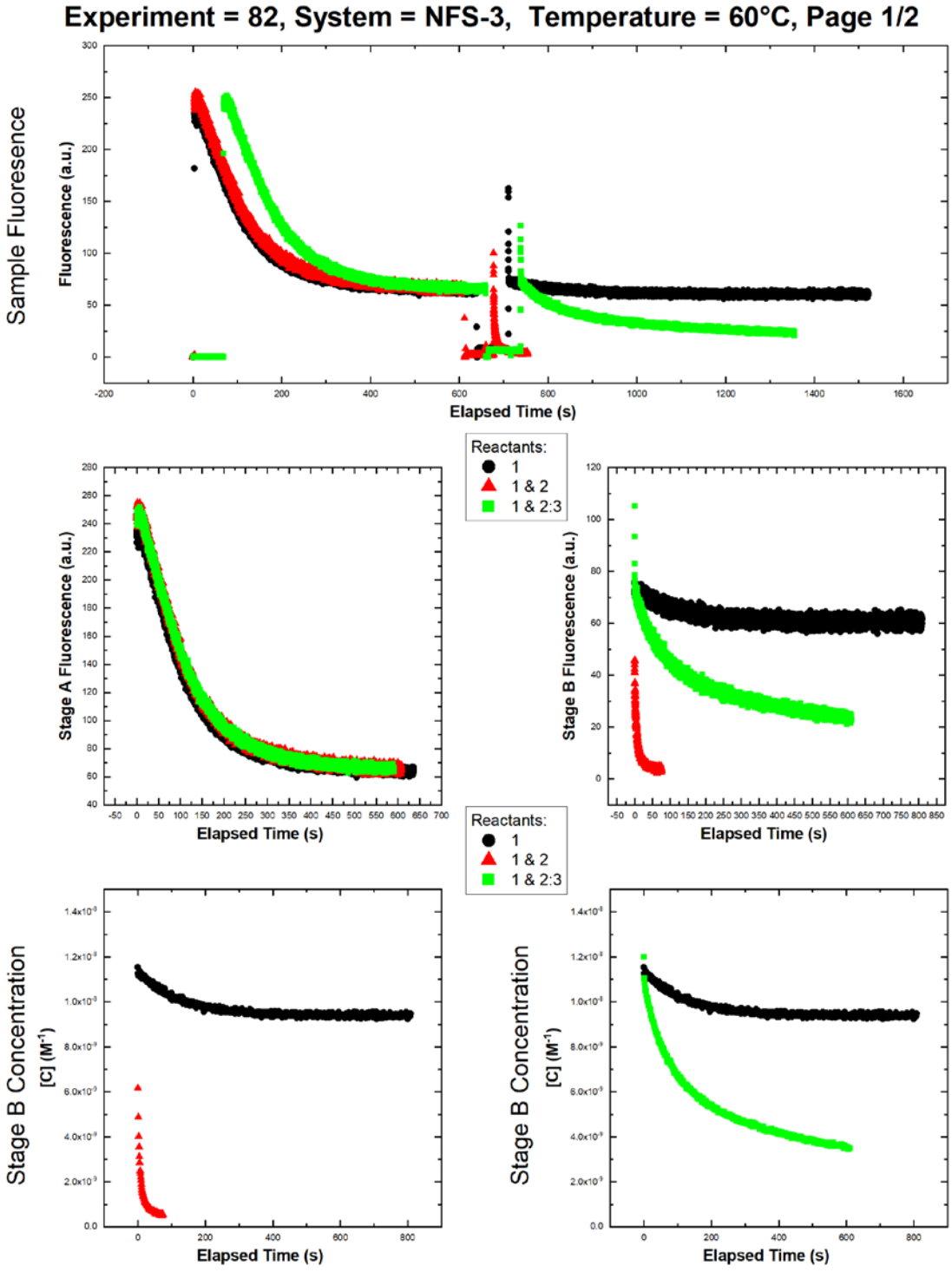

Experiment = 79, System = NFS-3, Temperature = 30°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

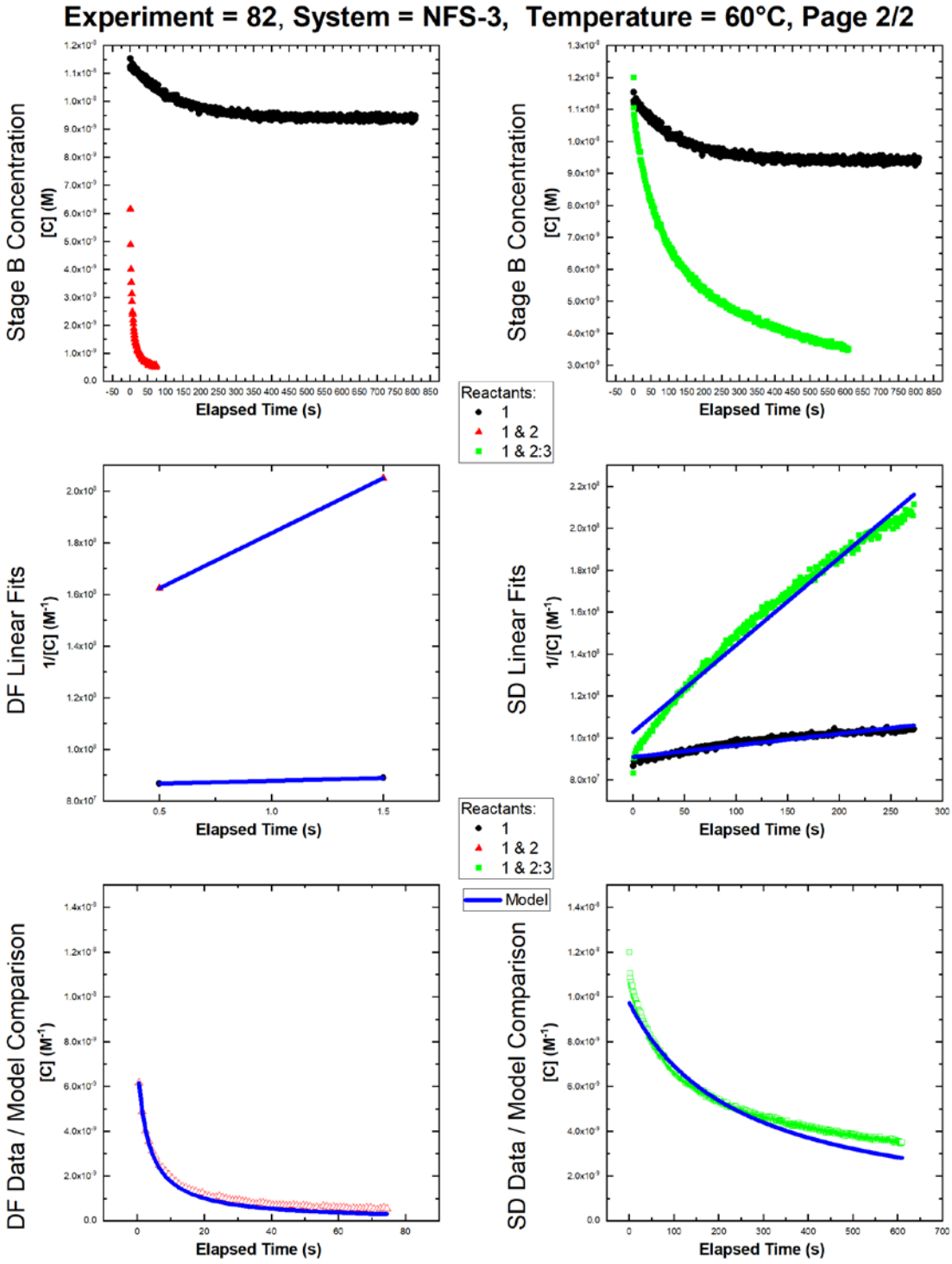

Experiment = 80, System = NFS-3, Temperature = 40°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)


Experiment = 80, System = NFS-3, Temperature = 40°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison


Experiment = 81, System = NFS-3, Temperature = 50°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)


Experiment = 81, System = NFS-3, Temperature = 50°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison


Experiment = 82, System = NFS-3, Temperature = 60°C, Page 1/2
Sample Fluoresence
Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage A Fluorescence (a.u.)
Elapsed Time (s)
Stage B Fluorescence (a.u.)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
Stage B Concentration
[C] (M-1)
Elapsed Time (s)
Stage B Concentration
[C] (M-1)
Elapsed Time (s)


Experiment = 82, System = NFS-3, Temperature = 60°C, Page 2/2
Stage B Concentration
[C] (M)
Elapsed Time (s)
Reactants:
1
1 & 2
1 & 2:3
DF Linear Fits
1/[C] (M-1)
SD Linear Fits
Model
DF Data / Model Comparison
SD Data / Model Comparison

APPENDIX B

New Sequences For Published Systems

Section B.1 The 10x10x10 DNA-Brick Structure

Table B.1 New sequences for the 10x10x10 DNA brick reported by Ke et al.[30]

Strand Name	Sequence (5'-3')
Strand-1	CAAATGCTCCGAAACCCGTGCGTGATCCTGAA
Strand-3	ATGTCAAAACCTCTCGCAGTCGTAAGTAAGTC
Strand-6	GGCTTCTGGGCGCGAGGCTTCCCGATTGACAC
Strand-7	TCCGCAACGTCTGTGTCCGAATATGTGAGGCC
Strand-8	TTATGCGTGGCTGTGGCTTGTTGAGACCCGAA
Strand-11	GTTGTGCAGAGGTCCATCCCGCTTAGGCTTAG
Strand-12	CGTGTCGAAAGGAAGTAGAGGTTGAGATGAGC
Strand-13	TGTCGCCGTATCTGTATGCACTGGTAAATGGT
Strand-16	ATCTTTGCGGAATGTTAATGACCTTTGTGTGT
Strand-17	AGGTCGAACATCAGTCTTGCTTGGAATCACAC
Strand-18	GCCCTCGAATAGTGCCCTTTAATAGTCTCATG
Strand-23	GCCCAGATCGTAAAGCCGGTGTATTCAAGCAT
Strand-25	ACAATAGGGCGGCAGGCGTCTCTCTTCGGGCA
Strand-27	GAGCTATTTGGTAAGTGCGGTTGGAAGTATCT
Strand-28	ATCACATCTTCCACTCGGTATTAAACATTCGG
Strand-29	ATATCACCCAGAGAAACACTACGTCATCCTTA
Strand-30	TTGGGAAAGCATCGACTCTAGCCGCACTGTAC
Strand-31	CTTAGCAATAGGGCGGCGTCATCTCGAAATAC
Strand-32	TGAAAGTTATGAGACAATGCCAAGAAGCGAGT
Strand-33	TTTGACACACCGATGGCACTGATTAGGCGAGG
Strand-34	TCTGCGAATACGACATAGTGAGCCTGGTTAAA
Strand-35	AGACCGTTGTTAGTAGAAATCCTTGCATGAAA
Strand-36	GCCAGAATATGCCGCCGTAGAAGAGTCGCCAA
Strand-37	TATGACAACTCCCTATTATCGGTGGTGAACGT
Strand-38	AAAGACTAACGAAAGCGGCAGACATAATTGAC
Strand-39	ATCTCCGGTAAACATGGTATTGACGCTATCTT
Strand-40	CCGTTTGCAGAAGCCAGGGAATCGCCCACTCC
Strand-41	TTACACATGCGTTCGTTATACTGATTTTCTCC
Strand-42	TGGCGTCCTAGTCTGAATCTGCTGTATAATCT
Strand-46	TCACCCTAGTTTTGTAAGAGTCCTGTCCATAA
Strand-48	ACCTCCTGGGCTGCTGATAGTTCCTCTCTGAT
Strand-49	TCCTGTACGATGGCGATATAACCTCCTAACTG
Strand-50	GTTGAGTGGTGAATGGGCCTGAATAGTCGGTT
Strand-51	GTAGTAGTATCCATTAAGTTCTGTATCATCAA
Strand-52	TAGGGACCAAATCATGCTGTGCGACCGAGGCG
Strand-53	CCACGGTAGCCACTGTCTGTATTATGAAGCGG
Strand-54	GCCTCCTAATGATGCTATACAAACAGTCGCGG
Strand-55	GCGTGGAGAAGCGTCGTTATACTAGCCTGCAT
Strand-56	CTAATTTGAGATCGTTTTGACTATGGGAGGCA
Strand-57	TGTCGGAGGCCACGCTTAGTCACGCAACCAAC

Strand Name	Sequence (5’-3’)
Strand-58	CGTATGGCGGGATTGTGCGGCGGACCTGGGCG
Strand-59	TTCATGTATCAAGGAGTTAAAAGTCTGTATCG
Strand-60	AGGTCACAAATCTCGGCCACAACACTACACTC
Strand-61	TCTCAGCATCTCTATATTTTAAGTACACCACA
Strand-62	TACACAGGCTTACCAGGGACGTTCTCTTATCG
Strand-63	GTTATGCAAAGGTTCAATCCAAATCGTCAAAA
Strand-64	CCAGTCGCTACGTTGTTATTTTCCGGGCCATG
Strand-67	TAGAAGCATCCCATCAACAAAGAACGTATTAG
Strand-68	TACATTTCGGGTGGCGGGTCGGCAGCTCCTCA
Strand-70	ACGAGTCCTTATAGCACGCCATCTTACCGCAG
Strand-72	CAAACCGTGGCAGTGATACCCTAGATGGAGAG
Strand-73	CTGCTAACGTAACAAGGAAAGAAATATGATGA
Strand-74	GATTCTGGTACACGATTTCATTGGGTTAGCAC
Strand-75	TACTGCGGATGGAGCGACGAACTTGGACCTTA
Strand-76	TACTGTCTAACGGGTCAAAGGGCTGCTCGGAT
Strand-77	GAAAGACAAACGCTAAGGTATCGGAATCAAGT
Strand-78	GATCGGTCCTTCTAGTTCGGCAGTACGGTTCA
Strand-79	CGGGACAAACGCTCTTAACGGGACTCTATGCC
Strand-80	GCTCGGACTCAGACGCGCTTAAAATGGTGCCT
Strand-81	CCGATTTATAGACTTCCGTTACAGAGCACAGC
Strand-82	TTAGCAACGGGTCACGGATGTCGCAGATGTCG
Strand-83	CACATCCATCCCAGACGTAATCGGTTGAGAGA
Strand-84	AGTCCATCGGCGCGGTAGTTCACGGTCAGTCA
Strand-85	GGCAGCTTTATCCGAACCCTCACTCCGTCAAC
Strand-86	ACACACGCGGGTCATGCCAAAGCCAGTCCAAT
Strand-87	TACAGTGATTCTTGATTATGGCTTGGACGGCT
Strand-91	CTGTGGGAGCGATAGGTAATGTGATGTTGCCG
Strand-93	ATGGATACAATGGGATGAGTCTTACAGCTTCT
Strand-94	TTCACAAGCCCTAAAGTACGTCGCCTATATGC
Strand-95	ATGCGGCCCTGTGAATTTAACGACTGATCCAA
Strand-96	CGCTGTTACCCAATACGAGAGATCCCTACGCT
Strand-97	TTAGAAAGTCATTATCAAAGGTACACAGCGAG
Strand-98	ATCAGTTAAGTCGAAGGAAGAGCCCGCCAGTT
Strand-99	GAATAGAACGCGCAGCTTAATGATCCCGGCCG
Strand-100	ACTATTGGCCCGCACGTTGTATCGCTAAGACT
Strand-101	TATATTCTAAAACTCGTGCGGTTCAGGCGGCG
Strand-102	CAGTTTTCCATGATCGCGCAAGAGAGTATAGT
Strand-103	CCTGGCTGGCTAGACAAACTCGCGGGCAATAC
Strand-104	TATATGTAATGAATGTGCTCCTCCAGTAGATG
Strand-105	AGTAAACCAGCGGAGTCACCCATCAGGCCATT
Strand-106	CCGCGACGTTCATCACTCAGGCTAATTTCCAC
Strand-107	AGGGACCACTGCTAGGGCGTTCGACAATTACC
Strand-108	ACATTGCGCTCATCACCTCGTCTAGTACCGGA

Strand Name	Sequence (5’-3’)
Strand-109	TGGCGCTTGGACTACTTCATGGGATACACGAC
Strand-112	GTCCCTAGAAATCCTCGTTGAATCCGCCGTCG
Strand-113	CCCATAACAGGCATACGGGAGGTGAGAATTTT
Strand-115	TGATCTGACGAGCGCCAATTCTGTTCTGAGCA
Strand-117	GATAGGTGACAAGGATAAAGGCTGAATACTGC
Strand-118	ACACAGCTATGCGGGTCCTTGATGCGTCCACT
Strand-119	TGAGATATTGAGTACGGTGTTTGAAGTTGTTA
Strand-120	GATACGTGAAAGAGATGGGTTGAGTCGAAATT
Strand-121	GAACATCGTCTGAAAGTAGGAACATTATCTGA
Strand-122	TGCAAATCAGGAGCTACTACTATATCTTAATG
Strand-123	ACAAGAACGGGAATACGACGAATCAAAGCAAG
Strand-124	GACAGACCCTACCCGACTTCACCCGTGTTGGA
Strand-125	AGCCGACCGTGGGATGTACTTCTGCCTGAGCC
Strand-126	TATCCTTGGTGTAGCGGTAAACTCAGAAAATG
Strand-127	TATAAGCAGTGTGTAACCTAGCCGCGTAAATG
Strand-128	CTAAACGGGTGCCTTCCTACAGAGATTCATTC
Strand-129	GCAGTTCGTGTCGCAAATACCTCGTCATCGTT
Strand-130	ATTCTATTATTACATTATATTAAACCTGCGTG
Strand-131	AGACTGCGACGATTACTAGAGACTTCATTTGT
Strand-132	CGACGACAAGGCCATACTGCTGCTAAACTGCG
Strand-136	GGTAAAGCCAGTCGATTCAGCGCCATAGATAA
Strand-138	TGTCGAGCTTGATTAGGGCAGTCCATTCCACC
Strand-139	CCGTCTAAGGGAAAGGAGTGCTGCCGTCATGG
Strand-140	ATGTTAGAGAATGCCGTATCAATCGTTTTTTA
Strand-141	AATCGAATCCAGTCAGTGTCTCTTTAGTGCCG
Strand-142	CATTTATCGCCCGCAGAGTCAGTTGGTCGTCC
Strand-143	TTCACGCCATTACGCAACGCTTACTTCAGACT
Strand-144	CAAGAGGCTAACGGATGATAAAACTAAAATTC
Strand-145	GCAATCGTAAAGGCGGGTCAAGCTAATCACTA
Strand-146	CAATCAAAAGGCGTCGTGTGGTCAGCGGAGGG
Strand-147	AGCGTATAGCCAATCCCTCGTTCACTCCACTG
Strand-148	AGTATCTCTTGTCTAATCCCTATCTGGTAAAA
Strand-149	CAGGGTAGAAACAAAATCCATGCTCGCTCGCT
Strand-150	CTAGCATATCTTACTGTCAACCACTGTCTTTA
Strand-151	GATCTGTTTAAGCATAGTCGGCGCGTGTCCCA
Strand-152	TTGAGGAGCCGAGCGCAGCCACCAGAGCCATG
Strand-153	GTATAGGTTAAGCCATGCGTTCAATCAGGAGT
Strand-154	TGAAATACCAAAACAGAACAAACTCGCGTCGT
Strand-157	TTGTCACTATATTAAACGGTCATGCAGCGGCC
Strand-158	CAACTATTTAAGCTATCGGCCATTCTACTCTA
Strand-160	GACCGTACATGGCTGGCGTTCACCCTCGATCC
Strand-162	GAAGGTCAGCTCCCGTCCCGCAAGGCCACACT
Strand-163	GGTCGTTGAGCATCCTTGAGTTGGCAGTAAAA

Strand Name	Sequence (5’-3’)
Strand-164	CCTTCGGGCGGCTCTCCTTCTCGGGTTGGCAC
Strand-165	GCCGCGCCTCAAACTATGTGTAAATCCGTGCA
Strand-166	TGAGTGATCGAAGTAATAAAGCAGCATGGTCT
Strand-167	TTGTAACCTCGTAGCCAGTCGGATAGAAGAAG
Strand-168	GGGCGGAAGCGATCTCTAAGCGACCACGGACT
Strand-169	CTCCAGGCTAAATCTTCCGCTTGTTGCAGGTG
Strand-170	TCTTCGCAAATTCGGGCCTTAACGTAGTCAGC
Strand-171	TATCCGCCAATGACACTTCGAGCTGCACGCCT
Strand-172	TGAAGGACTAAAAGCACTCTATGTTGGTACTG
Strand-173	ACGGAGGACTGGTGTTCATCGGACGGCATCGC
Strand-174	AAGTCCTATACAAGCCCGCCTGCAATGAGAAA
Strand-175	GTATCGTAGCGCCCGTGAATGACAGCCGTAGG
Strand-176	GAAACAGAGTCCAGCAAATCGCGGCAAGTACG
Strand-177	CACCTTCGTTATTATTCGGATTGCGGCCTGGG
Strand-181	CAAGGTGAGGACAACAGCGCCCGGATCTAAGT
Strand-183	TCTAAATGCTGACTGCGCCGGACCCTTGCCTT
Strand-184	GATAACGTCTAATGAGCTTCCCAGGCGCAGGC
Strand-185	TCTCATCTTAATCAGCTTGGCCTCGAATACTC
Strand-186	ACACTCACGGGTAAGCGGACATCAACGGGCAT
Strand-187	GTCAGGTCGCCGGATGCAGCCCACTTGTACGT
Strand-188	ATTGGTAGGTAAAAACAGCCGATTCTCATTCC
Strand-189	GCAAATTACGAACACGGACGAAACTGGAGCTG
Strand-190	AACGTGCTAGAGAGACTTATGCTTCCAGTGCG
Strand-191	TTGATGGTCCTTCTATTGTGCCCGATGCGCTG
Strand-192	CCGTACCGCTAGTGTGCTCGCTCGTGTTCTTG
Strand-193	GTTTCACTACTAAGATGCAGGACAATTCGTTC
Strand-194	TCCTATCCTTTCAGAGCATTGTTGGGCGAGAG
Strand-195	GATGCAACTGTGTCCCGTACCCGCGCCTGATG
Strand-196	AGAGATTTACGTGATGATTAACTATATGAATT
Strand-197	GTCCTCTATGAACGGCCCACTCTGTCTCTACG
Strand-198	ATTTAGGCGCAACCACTAGGCAAGGCTGACCC
Strand-199	TCACGTATGTAGACCACTTTAGCCCGATCCGC
Strand-202	TTGCCTTACTACTCGGGATATTTGCCTACATT
Strand-205	TCGCGCACGGGACTGGGAGGGTAGGACACACT
Strand-206	CCCAAGAACTCCCAGCGGCGATGTGATGCGTC
Strand-209	GCCCGAGTACCGCGAGCCCGAGAGTTGGACAC
Strand-210	GCTCATACAAGACCCTAAATACGAAGATACGA
Strand-211	TGGGTTTTGGATTTGGTTGGACATACTTCATA
Strand-214	ATAGGACCAGGCGGCTCGATGTTGTGTGAAGG
Strand-215	GCATACCTAACGGCTCTTGGACGAACTGGACT
Strand-216	GAGCACCCTGTCCACCAAAATGTCTACGCCGC
Strand-219	GACACGCCTGGACACTACATGATAACCCTACT
Strand-220	CATCGACCACTAGCTGTACGCGCATAAAAAGA

Strand Name	Sequence (5’-3’)
Strand-221	ACTGTGCCCTGGAAATAATGCTGCCCTGCGGT
Strand-227	TTTTTTTTGTTATAGGTTGAGCGATTTTTTTT
Strand-230	TTTTTTTTCTTACTCGGGACGCCATTTTTTTT
Strand-233	TTTTTTTTGGGAGGCTTTGTCATATTTTTTTT
Strand-236	TTTTTTTTCATATTGTAACTTTCATTTTTTTT
Strand-239	TTTTTTTTCCACCCGCAATAGCTCTTTTTTTT
Strand-240	CAGCGTGGTTTTTTTTTTTTTTTGGTCGGTG
Strand-241	CCTGCGAACGAAGCGAGAACTGAGACGGCCCA
Strand-243	TTTCCCTTGTGAGTGATTGAGTTGACCAGTTC
Strand-244	GCGAGACCCACGACCCACACTTGAGAGGACAG
Strand-245	TTTTTTTTGGCGACATTGCTTCTATTTTTTTT
Strand-246	ATACACCGTTTTTTTTTTTTTTTTGATGAGGG
Strand-247	ACGTAGTGATGCTTGATGTGACCTTATAGAGA
Strand-249	AGATGACGTGCCCGAACCTGTGTATGAACCTT
Strand-250	CCAACCGCGTATTTCGTCGAGGGCAGATCCAG
Strand-251	TTTTTTTTAGATACTTGCGACTGGTTTTTTTT
Strand-252	TTAATACCTTTTTTTTTTTTTTTTCTTCGCGC
Strand-253	GGCTCACTCCGAATGTCTCCACGCAACGATCT
Strand-255	TCTTCTACGTACAGTGCTCCGACAACAATCCC
Strand-256	CTTGGCATTTGGCGACCGGCGACAGTGAGACG
Strand-257	TTTTTTTTACTCGCTTTACATGAATTTTTTTT
Strand-258	AATCAGTGTTTTTTTTTTTTTTTCAATGCTT
Strand-259	GTCAATACCCTCGCCTCACTCAACTAATGGAT
Strand-261	TCAGTATATTTCATGCGGTCCCTAACAGTGGC
Strand-262	CACCGATAGGAGAAAAACGCATAATACGCTTG
Strand-263	TTTTTTTTACGTTCACTAGGAGGCTTTTTTTT
Strand-264	TGTCTGCCTTTTTTTTTTTTTTTGAATAACG
Strand-265	GAGCCGTGGTCAATTACTACTGCCTACAAAAC
Strand-267	GCGTTACAGGAGTGGGTTTATAGACAGCAGCC
Strand-268	CAGCAGATCGAAGAAGTTTGACATACAGGAAG
Strand-269	TTTTTTTTAGATTATAGTACAGGATTTTTTTT
Strand-270	TACCGACGTTTTTTTTTTTTTTTTGCTCTCTC
Strand-271	AGACCTAAACCCTAACGCGGCAAACATAAAAG
Strand-272	GTCAGCCAATGTGCCGTCCCTCCCGCGCAGTG
Strand-273	AAGTCCGATCCAATGCTGAGGCGTAGTAGGCG
Strand-274	AGGTTATATTTAGTTCGGCGGGCATTGAGCGT
Strand-275	TTTTTTTTCAGTTAGGCCACCGTCTTTTTTTT
Strand-276	ATTCAGGCTTTTTTTTTTTTTTTGCTTTCGT
Strand-277	AGGACTCTAACCGACTTGGATGTGACCGCGCC
Strand-278	TCGCACAGTTATGGACCCGGAGATTGGCTTCT
Strand-279	GGAACTATCGCCTCGGAAGCTGCCCATGACCC
Strand-280	GTTTGTATATCAGAGAATGTGTAATCAGACTA
Strand-281	TTTTTTTTCCGCGACTTCACTGTATTTTTTTT

Strand Name	Sequence (5’-3’)
Strand-282	TAGTATAATTTTTTTTTTTTTTTTCCATCGGT
Strand-283	ACAGAACTATGCAGGCGACCGATCAAGAGCGT
Strand-284	CGTGACTATTGATGATTTCGCAGACTACTAAC
Strand-285	TAATACAGGTTGGTTGGTCCGAGCGAAGTCTA
Strand-286	ACTTTTAACCGCTTCAATTCTGGCATAGGGAG
Strand-287	TTTTTTTTCGATACAGGTTGCTAATTTTTTTT
Strand-288	TGTTGTGGTTTTTTTTTTTTTTTGAGTGGAA
Strand-289	ATAGTCAAGAGTGTAGGTTAGCAGATCGTGTA
Strand-290	GAACGTCCTGCCTCCCGGTGATATGTCGATGC
Strand-291	TCCGCCGCCGATAAGACCGCAGTAGACCCGTT
Strand-292	GGAAAATACGCCCAGGTTGCTAAGTGTCTCAT
Strand-293	TTTTTTTTCATGGCCCTGTCTTTCTTTTTTTT
Strand-294	CGCTACGGTTTTTTTTTTTTTTTTGCTTTACG
Strand-295	ACTTAAAAACGCACTAGAAATGTACTGGGACG
Strand-296	CGGGAGGGTGTGGTGTGCCAACATCCTGCCGC
Strand-297	ATTTGGATCGCTACATGGACTCGTTAATATGG
Strand-298	TTCTTTGTTTTTGACGGCTTGCGGACTTACCA
Strand-299	TTTTTTTTCTAATACGACGGTTTGTTTTTTTT
Strand-300	CCGGCCTATTTTTTTTTTTTTTTTAACACATT
Strand-301	ATCACGCTGTCTGCTATCAAAATCTGGTCACG
Strand-302	CCGGCATCTGGCTCTCTATCGCATCTTGACAT
Strand-303	CACTCTCTGAAGGGAAACGACACGCATTAGGG
Strand-304	TGGGCACACTCGCACGGCTTTCGCTGATGGGA
Strand-305	TTTTTTTTACAATCAACTAGGGACTTTTTTTT
Strand-306	TGCCGACCTTTTTTTTTTTTTTTTCCGAGATT
Strand-307	CCAATGAATGAGGAGCGGTTTACTGTGATGAA
Strand-308	AGATGGCGGTGCTAACTGCTGAGACTGGTAAG
Strand-309	AGCCCTTTCTGCGGTATGGTCCCTGTGATGAG
Strand-310	CTAGGGTAATCCGAGCTGCATAACACAACGTA
Strand-311	TTTTTTTTCTCTCCATAAGCGCCATTTTTTTT
Strand-312	TTTCTTTCTTTTTTTTTTTTTTTTCGACGCTT
Strand-313	GTCCCGTTTCATCATACCAATAGTCGAGTTTT
Strand-314	AAGTTCGTGGCATAGACAAATTAGAGCGTGGC
Strand-315	CTGTAACGTAAGGTCCGAAAACTGTGTCTAGC
Strand-316	CCGATACCGCTGTGCTGCCATACGCTCCTTGA
Strand-317	TTTTTTTTACTTGATTTACATATATTTTTTTT
Strand-318	ACTGCCGATTTTTTTTTTTTTTTTCCATTCAC
Strand-319	CGTGAACTTGAACCGTGGCCGCATGTATTGGG
Strand-320	TTTTAAGCTGACTGACACTACTACCATGATTT
Strand-321	GGCTTTGGAGGCACCACTTTCTAACTTCGACT
Strand-322	GCGACATCATTGGACTTACCGTGGAGCATCAT
Strand-323	TTTTTTTTCGACATCTTTCTATTCTTTTTTTT
Strand-324	CCGATTACTTTTTTTTTTTTTTTTCTGGACCG

Strand Name	Sequence (5’-3’)
Strand-325	TCGTAATCTCTCTCAAACAGTCGGCCTATCGC
Strand-326	AGTGAGGGTGCGTCCGTAGGGTGAGAGTTTAG
Strand-327	ACAGGCAGGTTGACGGCTCCGGCGATCCCATT
Strand-328	AAGCCATATCATCAGCCAGGAGGTTCGCCATC
Strand-329	TTTTTTTTAGCCGTCCCTTGTGAATTTTTTTT
Strand-330	TACAGTCATTTTTTTTTTTTTTTTCGGCAACG
Strand-331	ATTATACGGGTCGCCTCAACGCAACATGGTGG
Strand-332	GCAGCAGCCTACAGCAGTCCACAGTTCTAAGT
Strand-333	GGGCAGGGCGTCTCCACCTCTCCGTCTAATGT
Strand-334	GCGACGTAGGTTTTAATCAATCCGAGTTCCAT
Strand-335	TTTTTTTTGCATATAGCTTCACGATTTTTTTT
Strand-336	GTCGTTAATTTTTTTTTTTTTTTTGTCTGGGA
Strand-337	TCACATTATTGGATCACCGTTTAGTTGCGACA
Strand-338	GTACCTTTCGGCAACAGATGGACTTTCGGATA
Strand-339	TAAGACTCCTCGCTGTAATAGAATGTAATCGT
Strand-340	ATCATTAAAGAAGCTGGCGTGTGTATCAAGAA
Strand-341	TTTTTTTTCGGCCGGGTGTCGTCGTTTTTTTT
Strand-342	CGATACAATTTTTTTTTTTTTTTTACTAGAAG
Strand-343	GATCTCTCAGTCTTAGGTTCTTGTTCGGGTAG
Strand-344	CTCTTGCGAGCGTAGGTTGTCCCGGCGTCTGA
Strand-345	GGCTCTTCACTATACTGGTCGGCTCGCTACAC
Strand-346	GGAGGAGCAACTGGCGTAAATCGGCGTGACCC
Strand-347	TTTTTTTTCATCTACTTGCTTATATTTTTTTT
Strand-348	GATGGGTGTTTTTTTTTTTTTTTTCTTGTTAC
Strand-349	GAACCGCAAATGGCCTAGCTGTGTCGTACTCA
Strand-350	TCGAACGCCGCCGCCTCCAGAATCCGCTCCAT
Strand-351	CGCGAGTTGGTAATTGCACGTATCCTTTCAGA
Strand-352	TCCCATGAGTATTGCCAGACAGTATTAGCGTT
Strand-353	TTTTTTTTGTCGTGTAGATTTGCATTTTTTTT
Strand-354	AGCACACGTTTTTTTTTTTTTTTTCGCCACCC
Strand-355	TAGCCTGAAGCACGGAGTTATGGGAAGACAAT
Strand-356	GAACCGCCGTGGAAATCTCGAACCTGCTATAA
Strand-357	TAGACGAGAGTATTTTTCAGATCATTGTGTGG
Strand-358	GATTCAACTCCGGTACTTTCGCTGTCACTGCC
Strand-359	TTTTTTTTCGACGGCGCACCTATCTTTTTTTT
Strand-360	TTACGAGCTTTTTTTTTTTTTTTTGGTTCTGG
Strand-361	TAGACTTTAGAATTAGCGGAACGGAACATACA
Strand-362	GGTTTTTCCTGGTTCCCGCACTCGACACATTG
Strand-363	CCGACTTATGACGCTGTGATGTAGGGCTGTAG
Strand-364	CTGGAACTACGCGCTAAGAACATTGAGGATTT
Strand-365	TTTTTTTTCTTTTATGAGTGACAATTTTTTTT
Strand-366	CACCTCCCTTTTTTTTTTTTTTTTACTCCGCT
Strand-367	TCAAACACAAAATTCTTATGCTAGTATGCTTA

Strand Name	Sequence (5’-3’)
Strand-368	ACAGAATTTAACAACTCGTCGCGGCCTAGCAG
Strand-369	TGTTCCTATGCTCAGACTCCTCAAATGGCTTA
Strand-370	CAGCCTTTTCAGATAACGCAATGTAGTAGTCC
Strand-371	TTTTTTTTGCAGTATTGTATTTCATTTTTTTT
Strand-372	CATCAAGGTTTTTTTTTTTTTTTTCGTGCGGG
Strand-373	GGGTGAAGAGTGGACGACGATTGCCGACGCCT
Strand-374	CTCAACCCTCCAACACAGAATATACGATCATG
Strand-375	GAGTTTACAATTTCGATATACGCTTTAGACAA
Strand-376	TATAGTAGCATTTTCTCAGCCAGGACATTCAT
Strand-377	TTTTTTTTCATTAAGACTACCCTGTTTTTTTT
Strand-378	GATTCGTCTTTTTTTTTTTTTTTTATTCACAG
Strand-379	CGAGGTATCTTGCTTTTCTAACATCTGACTGG
Strand-380	CAGAAGTAAACGATGATAACAGCGGATAATGA
Strand-381	AGTCTCTAGGCTCAGGGATAAATGTGCGTAAT
Strand-382	CGGCTAGGACAAATGATAACTGATGCTGCGCG
Strand-383	TTTTTTTTCATTTACGGCCTCTTGTTTTTTTT
Strand-384	CTCTGTAGTTTTTTTTTTTTTTTTCATGGTTT
Strand-385	GGGTAGCGGAATGAATAGAGGTAGATCGACTG
Strand-386	TTTAATATCTTCAACTTCCCACAGGTACATCT
Strand-387	TAATAACTCACGCAGGCTCAAGGGCTAATCAA
Strand-388	AGCAGCAGGGCAACGAGTATCCATCTTTAGGG
Strand-389	TTTTTTTTCGCAGTTTTTAGACGGTTTTTTTT
Strand-390	AGTATTAATTTTTTTTTTTTTTTTGTATGGAA
Strand-391	ACATGGTTCAAGACGTTTCCTGGCGACACACT
Strand-392	TTTTCAAGGTATTCGCTATCAACTGTAACTCT
Strand-393	GGCGGGATGCACCACGACTACAATTTTCTACG
Strand-394	GCAGCACTTTCCAGCGACGACTTCTTTCAGGG
Strand-395	TTTTTTTTCCATGACGATACTTAATTTTTTTT
Strand-396	GATTGATATTTTTTTTTTTTTTTTGAAGGCAC
Strand-397	GGCGCTGATAAAAAACTCCTCCGTGGCTTGTA
Strand-398	AACTGACTTTATCTATCGAACTGCAATGTAAT
Strand-399	GGACTGCCGGACGACCTACGATACTGCTGGAC
Strand-400	GTTTTATCGGTGGAATCGCAGTCTTATGGCCT
Strand-401	TTTTTTTTGAATTTTACGAAGGTGTTTTTTTT
Strand-402	AGCTTGACTTTTTTTTTTTTTTTTGTATTCCC
Strand-403	AAGAGACATAGTGATTTTCCGCCCAAGATTTA
Strand-404	TGAACGAGCGGCACTAGGTCTGTCCATCCCAC
Strand-405	GTAAGCGTCAGTGGAGTGCGAAGAGTGTCATT
Strand-406	AGCATGGAAGTCTGAACAAGGATATTACACAC
Strand-407	TTTTTTTTAGCGAGCGGTCCTTCATTTTTTTT
Strand-408	GTGGTTGATTTTTTTTTTTTTTTTACCCGCAT
Strand-409	TGACCACATAAAGACACAACGACCGAGAGCCG
Strand-410	TGGTGGCTCCCTCCGCATATCTCAATCTCTTT

Strand Name	Sequence (5’-3’)
Strand-411	GATAGGGACATGGCTCGGCGCGGCTTACTTCG
Strand-412	AGTTTGTTTTTTACCACGATGTTCTAGCTCCT
Strand-413	TTTTTTTTACGACGCGGGTTACAATTTTTTTT
Strand-414	GGTGCTTGTTTTTTTTTTTTTTTTGTATGCCT
Strand-415	GCGCCGACGTGCTTGAAATAGTTGGATGACCT
Strand-416	CTGAAACTTGGGACACACGGCATGGGCGCTCG
Strand-417	TTGAACGCGATATTAAGTACGGTCCCTGCTCT
Strand-418	CATGACCGACTCCTGACTGAACTAATCCTTGT
Strand-419	TTTTTTTTGGCCGCTGTGACCTTCTTTTTTTT
Strand-420	TAGAGCAATTTTTTTTTTTTTTTTCCCAGTGA
Strand-421	ATGATGATTAACCACCCACCGCTGAGGACCGT
Strand-422	CCGTAATGACCCTATATCGACAGACCGCGCTC
Strand-423	TAACTCTAGGCTGGCATGTGAAAACTCTGACT
Strand-424	GCCTGCTGACACCGCCGCTCCACTTTTAATAT
Strand-425	TTTTTTTTCACGATACTAAGGCAATTTTTTTT
Strand-426	AATGGCCGTTTTTTTTTTTTTTTTCAGTAAGA
Strand-427	CCGAGAAGTAGAGTAGGTTGCATCCATCACGT
Strand-428	GGTGAACGGTGCCAACAACAGATCGCGCTCGG
Strand-429	CTGCTTTAGGATCGAGTAGAGGACGTGGTTGC
Strand-430	CTTGCGGGAGACCATGACCTATACCTGTTTTG
Strand-431	TTTTTTTTAGTGTGGCATACGTGATTTTTTTT
Strand-432	CCAACTCATTTTTTTTTTTTTTTTCCGCCTTT
Strand-433	ACAAGCGGTTTTACTGAGCACGTTATAGAAGG
Strand-434	TTTACACACACCTGCATTTGATTGGGATTGGC
Strand-435	AGCTCGAATGCACGGACGGTACGGATCTTAGT
Strand-436	ATCCGACTAGGCGTGCGAGATACTTTTTGTTT
Strand-437	TTTTTTTTCTTCTTCTGGATAGGATTTTTTTT
Strand-438	GTCGCTTATTTTTTTTTTTTTTTTCGGCATTC
Strand-439	TGCAGGCGAGTCCGTGAGATGAGAGCTTACCC
Strand-440	CGTTAAGGTTTCTCATATTCGATTCTGCGGGC
Strand-441	CCGCGATTGCTGACTAGACCTGACGTTTTTAC
Strand-442	ACATAGAGCGTACTTGGGCGTGAAATCCGTTA
Strand-443	TTTTTTTTCAGTACCATAATTTGCTTTTTTTT
Strand-444	GTCCGATGTTTTTTTTTTTTTTTTAAGATTCA
Strand-445	TTTCCTGAGCGATGCCTAGCCGTATGTTGTCC
Strand-446	TGTCATTCTACTTTACGCTTTACCATATATTA
Strand-447	TTCTCGCACCTACGGCCTATCTAAGCAGTCAG
Strand-448	GCAATCCGGCTGGCTGGCTCGACACCTTTCCC
Strand-449	TTTTTTTTCCCAGGCCACGTTATCTTTTTTTT
Strand-450	ATTGTAGCTTTTTTTTTTTTTTTTGCACCGCG
Strand-452	GGAACTAATGCACCTCGATGACACGGGCTGAT
Strand-454	CTGGGAAGACTACTATGACTGGACGGGATTTG
Strand-455	TTTTTTTTGCCTGCGCCTTGGAACTTTTTTTT

Strand Name	Sequence (5’-3’)
Strand-456	GAGGCCAATTTTTTTTTTTTTTTTAACACCAG
Strand-458	GTGGGCTGACTTAGATTAGGACTTACGGGCGC
Strand-460	GTTTCGTCAAGGCAAGTCTGTTTCAATAATAA
Strand-461	TTTTTTTTCAGCTCCATTCCACGATTTTTTTT
Strand-462	AAGCATAATTTTTTTTTTTTTTTTGAGATCGC
Strand-464	CGAGCGAGATGCCCGTGCCTGGAGCCCGAATT
Strand-466	CAACAATGGGAATGAGGGCGGATATGCTTTTA
Strand-467	TTTTTTTTCTCTCGCCGTATCTCATTTTTTTT
Strand-468	GCGGGTACTTTTTTTTTTTTTTTTAGGATGCT
Strand-470	CAGAGTGGCAGCGCATCCCGAAGGTAGTTTGA
Strand-472	GGCTAAAGGAACGAATATCACTCAGGCTACGA
Strand-473	TTTTTTTTGCGGATCGTAACTGACTTTTTTTT
Strand-474	ACCCGCCGTTTTTTTTTTTTTTTTATAGCTTA
Strand-476	AAACTCCTAATTCATACTAAACTTCCAGCCAT
Strand-478	CAAATATCGGGTCAGCGGAAATGCACGGGAGC
Strand-479	TTTTTTTTAATGTAGGTAGCAATCTTTTTTTT
Strand-480	GGACTCACTTTTTTTTTTTTTTTTCATCGCTT
Strand-483	CAGCTACGTTTTTTTTTTTTTTTTGGGACACA
Strand-485	TGAAACGATATGAAGTGCCTAAATTGGTCTAC
Strand-486	ATAATCGGTTTTTTTTTTTTTTTTGTCTCTCT
Strand-488	TGGAATCTGCGGCGTAAGTGAAACCTCTGAAA
Strand-489	TGATATACTTTTTTTTTTTTTTTTGCTGATTA
Strand-491	GCTCCAAGACCGCAGGCTACCAATCGTGTTCG
Strand-492	GAGGGATGTTTTTTTTTTTTTTTTCTTAGTTG
Strand-494	TTCAGGTAGGTAGCTTCATTTAGACTCATTAG
Strand-1305	ACCTGTGACTTCGGAG
Strand-1307	AGCGTGCCCGGCCCTC
Strand-1309	CCACGCTGTCGCTTCG
Strand-1312	ACGCCTACGAGAGAGC
Strand-1319	TAGGCCGGTAGCAGAC
Strand-1322	TTTGCCGCCGTTGCCG
Strand-1329	GCTCGTAACTAATTCT
Strand-1332	TTGCGTTGTTCCATAC
Strand-1339	TTGCTCTAGGTGGTTA
Strand-1342	GCCAGGAACGCGGTGC
Strand-1349	GTGAGTCCTTATACCG
Strand-1357	GTACAATCGGTTTCGG
Strand-1361	ACAACAGTCTCGCGCC
Strand-1365	AGACACGTTGGACCTC
Strand-1369	TCCACTTCAACATTCC
Strand-1373	CCCAGCGCGAGGGCCG
Strand-1375	GCAAAGCCCGGTATAA
Strand-1379	CTCTCGGGTATCCAGG

Strand Name	Sequence (5'-3')
Strand-1383	CAACATCGCATCTAAT
Strand-1387	TATCATGTTGGGTTCC
Strand-1391	CTGACCGAGGCAACCG
Strand-1393	TACATTGGGCCTTGGG
Strand-2207	GATTGTACCGTTATTCTTTTTTTTGGGATGAG
Strand-2221	TTTTTTTTCTTCCTGTGATGCGAGCCTATAAC
Strand-2227	ACTGTTGTAAGCATTGTTTTTTTTGAGCCCTG
Strand-2241	TTTTTTTTCAAGCGTAGTCCCAGCCGAGTAAG
Strand-2247	ACGTGTCTGCGCGAAGTTTTTTTTGCTAGTTT
Strand-2261	TTTTTTTTCGTCTCACAATATACCAGCCTCCC
Strand-2267	GAAGTGGACCCTCATCTTTTTTTTGCTCATAG
Strand-2281	TTTTTTTTCTGGATCTCACACGTCACAATATG
Strand-2287	GCGCTGGGCACCGACCTTTTTTTTTAGATACCG
Strand-2293	TTTTTTTTCTGTCCTCGTGCAATAGCGGGTGG
Strand-3018	TCTGACACTTTTTTTTCGTAGCTGCCTGGATA
Strand-3036	GATTGCTAGTGTCGGTTCGTTTCATTTTTTTT
Strand-3038	ACCAACGATTTTTTTTCCGATTATATTAGATG
Strand-3056	GTCAGTTATACACATTAGATTCCATTTTTTTT
Strand-3058	TTTGCGAATTTTTTTTGTATATCAGGAACCCA
Strand-3076	TGAGATACATACTACCCTTGGAGCTTTTTTTT
Strand-3078	TTGGTTTGTTTTTTTTCATCCCTCCGGTTGCC
Strand-3096	TCGTGGAAGTGCTCTATACCTGAATTTTTTTT
Strand-3707	CGCCTCATAGCCCGTCTTTTTTTTCCGCACCT
Strand-3789	TTTTTTTTGACGCATCCTTGACCACCGAGTAG
Strand-3826	GAAATGGACGGAATAAACTTGCTAACACGGAAACTAATACACTTGCTG
Strand-3859	TCAAGTGTCCTGTCATAACCGAGAATCATCCTGCCGCGCCAAATCAAT
Strand-3861	ATGTTGGCAGGAGCTTTTCGCAGGTCCGTAACTTGTAATATCACTCAC
Strand-3862	CCGCAAGCTCTCGTCAAAGGGAAAGGGTCGTGGGTCTCGCATGTCGCC
Strand-3894	CGGGCTACCACTGCGCGGGAGGGAACGGATAACACGGCTCGGGCGGCT
Strand-3896	TCGCTCAAACGCTCAATGCCCGCCTCTAATGGTGTAACGCCTTCTTCG
Strand-3897	GGCAGTAGCGGTCCAGCGTCGGTAGTTAGGGTTTAGGTCTCGGCACAT
Strand-3898	TCTATAAACTAAACTCTGGCTGACGCATTGGATCGGACTTGAACTAAA
Strand-3929	ATGCGATATGGGCCGTCTCAGTTCAATGTGTTCCGTAGCGTAGTGCGT
Strand-3931	GCGAAAGCGAACTGGTCAACTCAAATGTCAAGCCCTCCCGATGTAGCG
Strand-3933	GGTTCGAGCGTCCCAGAGCGTGATGAGAGCCAGATGCCGGTTCCCTTC
Strand-3934	CAGCGAAACCATATTAAGAGAGTGCGTGCGAGTGTGCCCATTGATTGT
Strand-3966	ACGCCTCAACTTAGAACTGTGGACCTTTTATGGATTACGACGGACGCA
Strand-3968	GACGGTGGATGGAACTCGGATTGACGCCTACTCTGCCTGTGCTGATGA
Strand-3969	CCGACTGTAAACCATGTGACTGTAAGGCGACCCGTATAATTGCTGTAG
Strand-3970	CGCCGGAGAGATGTACGCTGCTGCTGGAGACGCCCTGCCCTTAAAACC
Strand-4001	CGAGTGCGCGTGACCAGATTTTGACCAGAACCCGTGTGCTTCCGTGCT
Strand-4003	AATGTTCTCCCTAATGCGTGTCGTCAATGTGTGGCGGTTCAAAATACT
Strand-4005	CATGCCGTATTGTCTTAAAGTCTAGGAACCAGGAAAAACCCAGCGTCA

Strand Name	Sequence (5'-3')
Strand-4006	TAGTTCAGCCACACAATAAGTCGGTAGCGCGTAGTTCCAGCATAAAAG
Strand-4038	CGGAGAGGAGAGTTACAGTTGATACCACCATGCGCTACCCAGTTGAAG
Strand-4040	TCGTGAAGCCCTGAAAGAAGTCGTACATTAGAAGTTATTATCGTTGCC
Strand-4041	CTACCTCTTGAATCTTTTAATACTACGTCTTGAACCATGTGCGAATAC
Strand-4042	CCCTTGAGTAATATATCTTGAAAACGTGGTGCATCCCGCCCGCTGGAA
Strand-4073	TCTGTCGATGTATGTTCCGTTCCGTCACTGGGCAAGCACCTCAAGCAC
Strand-4075	AGTGGAGCCTACAGCCCTACATCAGAGCGCGGAGTTTCAGTTAATATC
Strand-4077	AAGTTTAGAGGTCATCATCATCATTATAGGGTCATTACGGTGCCAGCC
Strand-4078	GCATTTCCAGAGCAGGTAGAGTTAGGCGGTGTCAGCAGGCGTATCGTG
Strand-4110	ATTGTAGTATCAGCCCGTGTCATCAGTGTGTCTCAGGAAAGTAAAGTA
Strand-4112	TTAAGTATCAAATCCCGTCCAGTCCGTAGAAATGCGAGAACAGCCAGC
Strand-4113	TACGGCTACAACTAAGGCTACAATGACGGGCTATGAGGCGGAGGTGCA
Strand-4114	TTAGATAGTGTGATAGTTAGTTCCGCCAGGTAGTAGCCCGATAGTAGT
Strand-4145	CCACAACTACGGTCCTCAGCGGTGAAGCGATGCGGCGGGTCTATACGA
Strand-4147	TGGTCAAGAGTCAGAGTTTTCACAAGCTAATAAGGAGTTTCAGATCAT
Strand-4149	GCTTGCTGCGACCTGCGGCTTTGCGACCCACCTAACCTGATGACTGCG
Strand-4182	AATAAGTTCCCAAGGCCCAATGTAAGGTGCGGTCGGTCAGCTCTCCCA
Strand-4184	GTTCCAAGAGTGTCGCGTGCCAGAGATCATGTTAATTCCCAAGCTACC
Strand-4185	TAGCAAGTCTCCGAAGTCACAGGTCTCATCCCGTAGGCGTTTATCCGT
Strand-4187	CTCGCATCCAGCAAGTGTATTAGTTTCCGTGTGTAGCCCGCCATTAGA
Strand-4188	CGATTCCCAGCCGCCCAGCATTTGTTATTCCGTCCATTTCCGAGAGGT
Strand-4189	ATATTCGGTTCAGGATCACGCACGCAGGGCTCTAGTCTTTCATGTTTA
Strand-4191	GCTGGGACGACTTACTTACGACTGGGCCTCACGCAAACGGACGAACGC
Strand-4192	AAGGATTTAAGATAGCCAGAAGCCACACAGACGTTGCGGACCACAGCC
Strand-4193	CAACCTCTGTGTCAATCGGGAAGCAAACTAGCGTGTCAAAATGTCGTA
Strand-4195	GGTATATTTTCGGGTCTCAACAAGGCTCATCTAACGGTCTGGCGGCAT
Strand-4196	CGGCTAGATTTAACCATGCACAACACTTCCTTTCGACACGTACAGATA
Strand-4197	CCAAGCAACTAAGCCTAAGCGGGACTATGAGCGATGTGATTTTCTCTG
Strand-4199	GACGTGTGACCATTTACCAGTGCAGTGTGATTTTTCCCAACCGCCCTA
Strand-4200	GAGAGACGTAAGGATGGCAAAGATGACTGATGTTCGACCTGGCACTAT
Strand-4201	GGCGCGGCACACACAAAGGTCATTCGGTATCTATCTGGGCAAGCTCCT
Strand-4203	TATTGCACCATGAGACTATTAAAGATTGATTTCCTATTGTTGACGAGA
Strand-4204	TATTACAAGTTACGGAGGCACGCTAGGATGATTCTCGGTTATGACAGG
Strand-4431	ACATCGCCCGCAGTCATCAGGTTAGGTGGGTCAGTTGTGGTATTAGCT
Strand-4433	TAGTTAATTCGTATAGGTGTCAGAGCAGGTCGCAGCAAGCCCAGTCCC
Strand-4434	CTTGCCTAATGATCTGGTGCGCGAGCTGGGAGTTCTTGGGACCGACAC
Strand-4436	ATGTCCAAAGTGTGTCCTACCCTCGTGTCCAAAAATCTCTGCCGTTCA
Strand-4438	CGGGCACACATCAGGCTCGTTGGTCTCGCGGTACTCGGGCAGGGTCTT
Strand-4439	TGTCCTGCCGTAGAGAGTATGAGCCCAAATCCAAAACCCAAATGTGTA
Strand-4441	GACATTTTTCGTATCTTCGTATTTCCTTCACAACCATCAACACACTAG
Strand-4443	TGATGTCCCGCACTGGTTCGCAAAAGCCGCCTGGTCCTATGAGCCGTT
Strand-4444	AATCGGCTCAAGAACAAGGTATGCGGTGGACAGGGTGCTCGGTAGTAT
Strand-4446	GCAGCATTAGTCCAGTTCGTCCAAAGTAGGGTGTGAGTGTCATCCGGC

Strand Name	**Sequence (5’-3’)**
Strand-4448	CCGGGCGCGAGTATTCCAAACCAAAGTGTCCAGGCGTGTCCAGCTAGT
Strand-4449	GGTCCGGCACGTACAAGGTCGATGATTTCCAGGGCACAGTTAGAGCAC
Strand-4451	GGGAATTATCTTTTTATGCGCGTATGGGAGAGTCACCTTGCTATCACA
Strand-4454	CGGGCTACTACCTGGCAACTTATTACATGATCTCTGGCACGCGACACT

Table B.2 Non-target structures present in the sequences for the 10x10x10 DNA-brick structure.[30]

	№ Structures	
Intramolecular	**Input[30]**	**Output (new)**
2 bp	15118	11775
3 bp	3197	67
4 bp	687	0
5 bp	143	0
6 bp	28	0
7 bp	6	0
8 bp	1	0

	№ Structures	
Intermolecular	**Input[30]**	**Output (new)**
2 bp	8632911	8620755
3 bp	2021642	2010751
4 bp	478168	465815
5 bp	114059	104831
6 bp	27783	22261
7 bp	5868	3846
8 bp	1612	339
9 bp	593	63
10 bp	160	2
11 bp	46	0
12 bp	21	0
13 bp	13	0
14 bp	12	0
15 bp	12	0
16 bp	12	0
17 bp	12	0
18 bp	3	0
19 bp	2	0
20 bp	2	0
21 bp	2	0
22 bp	2	0
23 bp	2	0
24 bp	2	0
25 bp	2	0

Section B.2 Four-Input OR Seesaw Gate System

Table B.3 New sequences for the "Four-Input OR gate" seesaw-gate based system published by Qian et al.[29]

Strand Name	Strand Sequence
w5,6	CAACCACAATAATCATCTCACCTCTAACCAACA
G5-b	TGAGATGTTGGTTAGAGGTGAGATG
w5,7	CAACTCTATAAATCATCTCACCTCTAACCAACA
Th2,5:5-t	CACCTCTAACCAACA
Th2,5:5-b	TGTTGGTTAGAGGTGAGATGTGTTGAGTTTT
w2,5	CACCTCTAACCAACATCTCAAAAACTCAACACA
G2-b	TGAGATGTGTTGAGTTTTTGAGATG
w1,2	CAAAAACTCAACACATCTCATTCTCCTACACCA
G1-b	TGAGATGGTGTAGGAGAATGAGATG
w1,10	CAAACAACTCTTACATCTCATTCTCCTACACCA
Th4,1:1-t	CATTCTCCTACACCA
Th4,1:1-b	TGGTGTAGGAGAATGAGATGGGTGTTTTAGT
w4,1	CATTCTCCTACACCATCTCAACTAAAACACCCA
G4-b	TGAGATGGGTGTTTTAGTTGAGATG
w3,2	CAAAAACTCAACACATCTCAATCCACACTATCA
G3-b	TGAGATGATAGTGTGGATTGAGATG
w3,11	CACTTACAAACTACATCTCAATCCACACTATCA
Th12,3:3-t	CAATCCACACTATCA
Th12,3:3-b	TGATAGTGTGGATTGAGATGAGGATTTTGTG
w12,3	CAATCCACACTATCATCTCACACAAAATCCTCA
G12-b	TGAGATGAGGATTTTGTGTGAGATG
G8-b	TGAGATGTTATTTGGTGATGAGATG
w8,15	CAAATCTACTCTACATCTCATCACCAAATAACA
Th16,8:8-t	CATCACCAAATAACA
Th16,8:8-b	TGTTATTTGGTGATGAGATGAAGATTAGGTT
w16,8	CATCACCAAATAACATCTCAAACCTAATCTTCA
G16-b	TGAGATGAAGATTAGGTTTGAGATG
G17-b	TGAGATGGTAGAAGTTTATGAGATG
w17,19	CAACAACTCTCTACATCTCATAAACTTCTACCA
Th20,17:17-t	CATAAACTTCTACCA
Th20,17:17-b	TGGTAGAAGTTTATGAGATGGAGTTAGTATG
w20,17	CATAAACTTCTACCATCTCACATACTAACTCCA
G20-b	TGAGATGGAGTTAGTATGTGAGATG
w21,20	CACATACTAACTCCATCTCACTCTAAACAAACA
w22,20	CACATACTAACTCCATCTCACTTTCATTTCACA
w18,16	CAAACCTAATCTTCATCTCACTACTCTATATCA
w9,4	CAACTAAAACACCCATCTCACTACAAACAATCA
w13,12	CACACAAAATCCTCATCTCACTCTCTACAAACA
w14,12	CACACAAAATCCTCATCTCACTCTATCTAAACA

Strand Name	Strand Sequence
w8,2	CAAAAACTCAACACATCTCATCACCAAATAACA
w17,2	CAAAAACTCAACACATCTCATAAACTTCTACCA
w23,4	CAACTAAAACACCCATCTCACTCTCTACAATCA
w24,16	CAAACCTAATCTTCATCTCACTCTCTCTATACA
Rep6-t	CAACCACAATAATCA
Rep6-b	TGATTATTGTGGTTGAGATG

Table B.4 Non-target structures present in the "Four-Input OR" seesaw-gate based system.

	№ Interferences	
Intramolecular	**Input**[29]	**Output (new)**
2 bp	308	218
3 bp	10	0
4 bp	1	0

	№ Interferences	
Intermolecular	**Input**[29]	**Output (new)**
2 bp	46094	42771
3 bp	13915	12998
4 bp	4375	3422
5 bp	1438	680
6 bp	635	204
7 bp	319	182
8 bp	283	182
9 bp	85	11
10 bp	20	0
11 bp	16	0
12 bp	15	0
13 bp	14	0
14 bp	14	0
15 bp	14	0
16 bp	14	0
17 bp	14	0
18 bp	14	0
19 bp	14	0
20 bp	9	0
21 bp	6	0
22 bp	2	0

Section B.3 Autocatalytic Four-Arm-Junction Network

Table B.5 New sequences for the autocatalytic-four-arm-junction system published by Kotani et al.[65]

Strand Name	Strand Sequence (5’ -> 3’)
A1x	GGTGTAGCGTGCGTAGAGATGCGTGTGTC-AAGGTAAGCGGTAGGTTCGTCCAAAGGTG
fA2au	CGCTTACCTTGGACCGGACCTGGGCTGAC-CTGAACACACGCATCTCTACGCACGCTACACCTC
fA3au	CAGCAGTCCCATTCCCAGCCAGTCAGGTC-AGCCCAGGTCCGGTCTTACACACGCATCTCTACGCACGC
fA4au	GACGAACCTACCGCTTACCTTGCTGGCTG-GGAATGGGACTGCTGCTACTGCTCTCACTCA
B1x	CACCTTTGGACGAACCTACCGCTTACCTT-GACACACGCATCTCTACGCACGC
fB2au	GAGGTGTAGCGTGCGTAGAGATGCGTGT-GTTCAGGTCAGCCCAGGTCCGGTCCAAGGTAAGCG
fB3au	CACCTTTGGACGAACCTACGACGAACCTA-CGACCGGACCTGGGCTGACCTGACTGGCTGGGAATGGGACTGCTG
fB4au	TTCTCCATCCACATCATTCAGCAGTCCCAT-TCCCAGCCAGCAAGGTAAGCGGTAGGTTCGTCCAAAGG
dye	CTTTCTCCATCCACATCACTACTG
quencher	TGAGTGAGAGCAGTAGTGATGTGGATGGAGAAAG

Table B.6 Non-target structures in the autocatalytic-four-arm-junction system published by Kotani et al.

	№ Interferences	
	Input[65]	**Output (new)**
Intramolecular		
2 bp	779	573
3 bp	157	0
4 bp	26	0
5 bp	1	0
Intermolecular	**Input**[65]	**Output (new)**
2 bp	10395	10272
3 bp	2516	2269
4 bp	622	666
5 bp	137	133
6 bp	46	16
7 bp	13	0
8 bp	11	0
9 bp	11	0
10 bp	11	0
11 bp	11	0
12 bp	11	0
13 bp	7	0
14 bp	3	0
15 bp	3	0
16 bp	3	0
17 bp	3	0
18 bp	1	0
19 bp	1	0
20 bp	1	0
21 bp	1	0
22 bp	1	0
23 bp	1	0
24 bp	1	0
25 bp	1	0
26 bp	1	0
27 bp	1	0
28 bp	1	0
29 bp	1	0
30 bp	1	0
31 bp	1	0
32 bp	1	0
33 bp	1	0

34 bp	1	0
35 bp	1	0
36 bp	1	0
37 bp	1	0
38 bp	1	0
39 bp	1	0
40 bp	1	0
41 bp	1	0
42 bp	1	0
43 bp	1	0
44 bp	1	0
45 bp	1	0

Section B.4 Autocatalytic Entropy-Driven Network

Table B.7 New sequences for the autocatalytic system published by Zhang et al.[25]

Strand Name	Strand Sequence (5' - 3')
Catalyst	ACCATTACTACACGCTTCCACTTATTCAGACGAC
Signal	TCTCTATCAACAAACTCCTCACCATTACTACACGCT
Backbone	AGTGGAAGCGTGTAGTAATGGTGAGGGTCGTCTGAATAAGTGGAAGCGTG
Fuel	CACGCTTCCACTTATTCAGACGACCCTCACCATTACTACACGCT
Dye	TCTCTATCAACAAACTCCTC
Quencher	AGCGTGTAGTAATGGTGAGGAGTTTGTTGATAGAGA

Table B.8 Non-target structures in the autocatalytic system published by Zhang et al.[25]

	№ Interferences	
Intramolecular	**Input[25]**	**Output (new)**
2 bp	146	67
3 bp	22	0
4 bp	3	0
Intermolecular	**Input[25]**	**Output (new)**
2 bp	1542	1372
3 bp	328	206
4 bp	77	0

REFERENCES

1 Glansdorff, N., Xu, Y. & Labedan, B. The Last Universal Common Ancestor: emergence, constitution and genetic legacy of an elusive forerunner. *Biol Direct* **3**, doi:Artn 29 10.1186/1745-6150-3-29 (2008).

2 Zheng, J. *et al.* From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **461**, 74-77, doi:10.1038/nature08274 (2009).

3 Bathe, M. & Rothemund, P. W. K. DNA nanotechnology: A foundation for programmable nanoscale materials. *MRS Bulletin* **42**, 882-888, doi:10.1557/mrs.2017.279 (2017).

4 Bianconi, E. *et al.* An estimation of the number of cells in the human body. *Ann Hum Biol* **40**, 463-471, doi:10.3109/03014460.2013.807878 (2013).

5 Sender, R., Fuchs, S. & Milo, R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *Plos Biology* **14**, doi:ARTN e1002533 10.1371/journal.pbio.1002533 (2016).

6 Chen, Y. J., Groves, B., Muscat, R. A. & Seelig, G. DNA nanotechnology from the test tube to the cell. *Nat Nanotechnol* **10**, 748-760, doi:10.1038/nnano.2015.195 (2015).

7 Jung, C. & Ellington, A. D. Diagnostic applications of nucleic acid circuits. *Acc Chem Res* **47**, 1825-1835, doi:10.1021/ar500059c (2014).

8 Kumar, V. *et al.* DNA nanotechnology for cancer therapy. *Theranostics* **6**, 710-725, doi:10.7150/thno.14203 (2016).

9 Chhabra, R., Sharma, J., Liu, Y., Rinker, S. & Yan, H. DNA self-assembly for nanomedicine. *Adv Drug Deliv Rev* **62**, 617-625, doi:10.1016/j.addr.2010.03.005 (2010).

10 Church, G. M., Gao, Y. & Kosuri, S. Next-Generation Digital Information Storage in DNA. *Science* **337**, 1628-1628, doi:10.1126/science.1226355 (2012).

11 Goldman, N. *et al.* Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature* **494**, 77-80, doi:10.1038/nature11875 (2013).

12 Zhirnov, V., Zadegan, R. M., Sandhu, G. S., Church, G. M. & Hughes, W. L. Nucleic acid memory. *Nat Mater* **15**, 366-370, doi:10.1038/nmat4594 (2016).

13 Watson, J. D. & Crick, F. H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**, 737-738 (1953).

14 Franklin, R. E. & Gosling, R. G. Molecular Configuration in Sodium Thymonucleate. *Nature* **171**, 740-741, doi:DOI 10.1038/171740a0 (1953).

15 Zhang, D. Y. in *Lect Notes Comput Sc.*

16 Watson, J. D. *Molecular biology of the gene*. 6th edn, (Pearson/Benjamin Cummings; Cold Spring Harbor Laboratory Press, 2008).

17 Ghosh, A. & Bansal, M. A glossary of DNA structures from A to Z. *Acta Crystallographica Section D-Structural Biology* **59**, 620-626, doi:10.1107/S0907444903003251 (2003).

18 Seeman, N. C. Nucleic acid junctions and lattices. *J Theor Biol* **99**, 237-247 (1982).

19 Chen, J. H. & Seeman, N. C. Synthesis from DNA of a molecule with the connectivity of a cube. *Nature* **350**, 631-633, doi:10.1038/350631a0 (1991).

20 Winfree, E., Liu, F., Wenzler, L. A. & Seeman, N. C. Design and self-assembly of two-dimensional DNA crystals. *Nature* **394**, 539-544, doi:10.1038/28998 (1998).

21 Yin, P., Yan, H., Daniell, X. G., Turberfield, A. J. & Reif, J. H. A unidirectional DNA walker that moves autonomously along a track. *Angew Chem Int Ed Engl* **43**, 4906-4911, doi:10.1002/anie.200460522 (2004).

22 Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297-302, doi:10.1038/nature04586 (2006).

23 Seelig, G., Soloveichik, D., Zhang, D. Y. & Winfree, E. Enzyme-free nucleic acid logic circuits. *Science* **314**, 1585-1588, doi:10.1126/science.1132493 (2006).

24 Seelig, G., Yurke, B. & Winfree, E. Catalyzed relaxation of a metastable DNA fuel. *J Am Chem Soc* **128**, 12211-12220, doi:10.1021/ja0635635 (2006).

25 Zhang, D. Y., Turberfield, A. J., Yurke, B. & Winfree, E. Engineering entropy-driven reactions and networks catalyzed by DNA. *Science* **318**, 1121-1125, doi:10.1126/science.1148532 (2007).

26 Yin, P., Choi, H. M., Calvert, C. R. & Pierce, N. A. Programming biomolecular self-assembly pathways. *Nature* **451**, 318-322, doi:10.1038/nature06451 (2008).

27 Zhang, D. Y. & Winfree, E. Control of DNA strand displacement kinetics using toehold exchange. *J Am Chem Soc* **131**, 17303-17314, doi:10.1021/ja906987s (2009).

28 Soloveichik, D., Seelig, G. & Winfree, E. DNA as a universal substrate for chemical kinetics. *Proc Natl Acad Sci U S A* **107**, 5393-5398, doi:10.1073/pnas.0909380107 (2010).

29 Qian, L. & Winfree, E. Scaling up digital circuit computation with DNA strand displacement cascades. *Science* **332**, 1196-1201, doi:10.1126/science.1200520 (2011).

30 Ke, Y. G., Ong, L. L., Shih, W. M. & Yin, P. Three-dimensional structures self-assembled from DNA bricks. *Science* **338**, 1177-1183, doi:10.1126/science.1227268 (2012).

31 Zhang, D. Y., Hariadi, R. F., Choi, H. M. & Winfree, E. Integrating DNA strand-displacement circuitry with DNA tile self-assembly. *Nat Commun* **4**, 1965, doi:10.1038/ncomms2965 (2013).

32 Li, B. L., Ellington, A. D. & Chen, X. Rational, modular adaptation of enzyme-free DNA circuits to multiple detection methods. *Nucleic Acids Research* **39**, doi:ARTN e110 10.1093/nar/gkr504 (2011).

33 Ong, L. L. *et al.* Programmable self-assembly of three-dimensional nanostructures from 10,000 unique components. *Nature* **552**, 72-77, doi:10.1038/nature24648 (2017).

34 Dirks, R. M., Lin, M., Winfree, E. & Pierce, N. A. Paradigms for computational nucleic acid design. *Nucleic Acids Res* **32**, 1392-1403, doi:10.1093/nar/gkh291 (2004).

35 Woo, S. & Rothemund, P. W. K. Programmable molecular recognition based on the geometry of DNA nanostructures. *Nature Chemistry* **3**, 620-627, doi:10.1038/Nchem.1070 (2011).

36 Zhang, J. X. *et al.* Predicting DNA hybridization kinetics from sequence. *Nat Chem* **10**, 91-98, doi:10.1038/nchem.2877 (2018).

37 Hata, H., Kitajima, T. & Suyama, A. Influence of thermodynamically unfavorable secondary structures on DNA hybridization kinetics. *Nucleic Acids Res* **46**, 782-791, doi:10.1093/nar/gkx1171 (2018).

38 Yurke, B., Turberfield, A. J., Mills, A. P., Jr., Simmel, F. C. & Neumann, J. L. A DNA-fuelled molecular machine made of DNA. *Nature* **406**, 605-608, doi:10.1038/35020524 (2000).

39 Seeman, N. C. Nanomaterials based on DNA. *Annu Rev Biochem* **79**, 65-87, doi:10.1146/annurev-biochem-060308-102244 (2010).

40 Seeman, N. C. De novo design of sequences for nucleic acid structural engineering. *J Biomol Struct Dyn* **8**, 573-581, doi:10.1080/07391102.1990.10507829 (1990).

41 Feldkamp, U., Saghafi, S., Banzhaf, W. & Rauhe, H. in *DNA Computing*. 23-32 (Springer Berlin Heidelberg).

42 Kick, A., Bonsch, M. & Mertig, M. EGNAS: an exhaustive DNA sequence design algorithm. *BMC Bioinformatics* **13**, 138, doi:10.1186/1471-2105-13-138 (2012).

43 Tanaka, F., Kameda, A., Yamamoto, M. & Ohuchi, A. Design of nucleic acid sequences for DNA computing based on a thermodynamic approach. *Nucleic Acids Research* **33**, 903-911, doi:10.1093/nar/gki235 (2005).

44 Birac, J. J., Sherman, W. B., Kopatsch, J., Constantinou, P. E. & Seeman, N. C. Architecture with GIDEON, a program for design in structural DNA nanotechnology. *J Mol Graph Model* **25**, 470-480, doi:10.1016/j.jmgm.2006.03.005 (2006).

45 Andersen, E. S. *et al.* DNA origami design of dolphin-shaped structures with flexible tails. *Acs Nano* **2**, 1213-1218, doi:10.1021/nn800215j (2008).

46 Douglas, S. M. *et al.* Rapid prototyping of 3D DNA-origami shapes with caDNAno. *Nucleic Acids Research* **37**, 5001-5006, doi:10.1093/nar/gkp436 (2009).

47 Williams, S. *et al.* 90-101 (Springer Berlin Heidelberg).

48 Zhu, J., Wei, B., Yuan, Y. & Mi, Y. UNIQUIMER 3D, a software system for structural DNA nanotechnology design, analysis and evaluation. *Nucleic Acids Res* **37**, 2164-2175, doi:10.1093/nar/gkp005 (2009).

49 Andersen, E. S. Prediction and design of DNA and RNA structures. *N Biotechnol* **27**, 184-193, doi:10.1016/j.nbt.2010.02.012 (2010).

50 Zadeh, J. N. *et al.* NUPACK: Analysis and design of nucleic acid systems. *J Comput Chem* **32**, 170-173, doi:10.1002/jcc.21596 (2011).

51 Beliveau, B. J. *et al.* OligoMiner provides a rapid, flexible environment for the design of genome-scale oligonucleotide in situ hybridization probes. *Proc Natl Acad Sci USA* **115**, E2183-E2192, doi:10.1073/pnas.1714530115 (2018).

52 Wolfe, B. R., Porubsky, N. J., Zadeh, J. N., Dirks, R. M. & Pierce, N. A. Constrained multistate sequence design for nucleic acid reaction pathway engineering. *J Am Chem Soc* **139**, 3134-3144, doi:10.1021/jacs.6b12693 (2017).

53 Wang, J. S. & Zhang, D. Y. Simulation-guided DNA probe design for consistently ultraspecific hybridization. *Nature Chemistry* **7**, 545-553, doi:10.1038/Nchem.2266 (2015).

54 Olson, X. *et al.* Availability: A Metric for Nucleic Acid Strand Displacement Systems. *Acs Synth Biol* **6**, 84-93, doi:10.1021/acssynbio.5b00231 (2017).

55 Wetmur, J. G. & Davidson, N. Kinetics of Renaturation of DNA. *J Mol Biol* **31**, 349-&, doi:Doi 10.1016/0022-2836(68)90414-2 (1968).

56 Wallace, M. I., Ying, L. M., Balasubramanian, S. & Klenerman, D. Non-Arrhenius kinetics for the loop closure of a DNA hairpin. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 5584-5589, doi:DOI 10.1073/pnas.101523498 (2001).

57 Rauzan, B. *et al.* Kinetics and Thermodynamics of DNA, RNA, and Hybrid Duplex Formation. *Biochemistry* **52**, 765-772, doi:10.1021/bi3013005 (2013).

58 Sikora, J. R., Rauzan, B., Stegemann, R. & Deckert, A. Modeling Stopped-Flow Data for Nucleic Acid Duplex Formation Reactions: The Importance of Off-Path Intermediates. *Journal of Physical Chemistry B* **117**, 8966-8976, doi:10.1021/jp404550a (2013).

59 Sanstead, P. J. & Tokmakoff, A. Direct Observation of Activated Kinetics and Downhill Dynamics in DNA Dehybridization. *Journal of Physical Chemistry B* **122**, 3088-3100, doi:10.1021/acs.jpcb.8b01445 (2018).

60 Chen, C. L., Wang, W. J., Wang, Z., Wei, F. & Zhao, X. S. Influence of secondary structure on kinetics and reaction mechanism of DNA hybridization. *Nucleic Acids Research* **35**, 2875-2884, doi:10.1093/nar/gkm177 (2007).

61 Franck, J. & Rabinowitsch, E. Some remarks about free radicals and the photochemistry of solutions. *T Faraday Soc* **30**, 0120-0130, doi:DOI 10.1039/tf9343000120 (1934).

62 Turberfield, A. J. *et al.* DNA fuel for free-running nanomachines. *Phys Rev Lett* **90**, doi:ARTN 118102 10.1103/PhysRevLett.90.118102 (2003).

63 Wyer, J. A., Kristensen, M. B., Jones, N. C., Hoffmann, S. V. & Nielsen, S. B. Kinetics of DNA duplex formation: A-tracts versus AT-tracts. *Phys Chem Chem Phys* **16**, 18827-18839, doi:10.1039/c4cp02252a (2014).

64 Srinivas, N. *et al.* On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Research* **41**, 10641-10658, doi:10.1093/nar/gkt801 (2013).

65 Kotani, S. & Hughes, W. L. Multi-Arm Junctions for Dynamic DNA Nanotechnology. *J Am Chem Soc* **139**, 6363-6368, doi:10.1021/jacs.7b00530 (2017).