

# Combining MD simulations with sPRE calculations for protein structure prediction

Kevin Sawade<sup>a</sup>

**Keywords:** Keywords; L<sup>A</sup>T<sub>E</sub>X 2<sub>E</sub>; Stat

**Abstract:** MD simulations are a valuable tool to gain insights into protein structure and dynamics. The atomistic resolution paired with the time resolution of pico seconds have enabled looks into the dynamic behavior of protein folding and unfolding. Compared to other methods which yield time-averaged results, MD can be used as a tool to identify sub-states and predict transition paths between such states. However, reaching such conclusions requires large datasets, that can be different to obtain. System sizes and their number of atoms scale with  $r^3$ . The folding of some proteins might occur on such long timescales that all-atom simulations might not be feasible. In this manuscript we present a method to combine a limited set of all-atom MD data with a large dataset of coarse-grained MD data and solvent paramagnetic resonance enhancement (sPRE) spectroscopy. We applied this method to three differently ubiquitin chains. Copyright © 2018 John Wiley & Sons, Ltd.

## 1. Introduction

### 1.1. Available data

**CG data** MD data at coarse-grained resolution was taken from (?)

## 2. Getting Started

The **WileySTAT-V1** class file should run on any standard L<sup>A</sup>T<sub>E</sub>X 2<sub>E</sub> installation. If any of the fonts, class files or packages it requires are missing from your installation, they can be found on the *T<sub>E</sub>X Live* or from CTAN.

---

<sup>a</sup>Universität Konstanz, Universitätsstrasse 10, 78464 Konstanz, Germany  
\*Email: an@firstauthor

This is article note. Processes play a keyrole in the kinetics of many microstructural changes that occur during processing of metals, alloys, ceramics, semiconductors, glasses, and polymers.

### 3. Article Usage

*Stat* is published using proprietary fonts. A reasonable match can be achieved by using the `cmbright` option as `\documentclass[cmbright]{WileySTAT-V1}`. If for any reason you have a problem using the CM Bright fonts you can easily resort to Computer Modern fonts by removing the `cmbright` option.

Use `doublespace` option as `\documentclass[cmbright,doublespace]{WileySTAT-V1}` for extra line spread.

- (i) In `\runninghead` it should contain the name of author/s (verso) and short title (recto). If there are three or more authors, list down the surname of first author (with initial) followed by '*et al.*'.
- (ii) Authors and Addresses: In the `\author{}` tag, put your names in full, the firstname should not be an initial. For authors with affiliations within the US, please indicate your complete address including the state and zip code.
- (iii) Note the use of `\affil` and `\affilnum` to link names and addresses. For authors that share in the same address, they should have the same link letter.
- (iv) The author for correspondence (via email) is marked by `\corrauth` and `\correemail`. The latter is used to give the corresponding author's email, to be printed as a footnote prefaced by 'Email:'.
- (v) The abstract should be capable of standing by itself, in the absence of the body of the article and of the bibliography. Therefore, it must not contain any reference citations.
- (vi) Keywords are separated by semicolons.

### 4. Heading 1: Coding text content

This chapter will concentrate on bulk diffusion in solid metals and alloys. Most of the solid elements are metals. Furthermore, diffusion properties and atomic mechanisms of diffusion have most thoroughly been investigated in metallic solids.<sup>1</sup>

#### 4.1. Heading 2: Coding text content

This chapter will concentrate on bulk diffusion in solid metals and alloys. Most of the solid elements are metals. Furthermore, diffusion properties and atomic mechanisms of diffusion have most thoroughly been investigated in metallic solids.

##### 4.1.1. Heading 3: Coding text content

This chapter will concentrate on bulk diffusion in solid metals and alloys. Most of the solid elements are metals. Furthermore, diffusion properties and atomic mechanisms of diffusion have most thoroughly been investigated in metallic solids.

**Heading 4: Coding text content.** This chapter will concentrate on bulk diffusion in solid metals and alloys. Most of the solid elements are metals. Furthermore, diffusion properties and atomic mechanisms of diffusion have most thoroughly been investigated in metallic solids.

<sup>1</sup>This is footnote text. For anisotropic media and non-cubic crystalline solids  $\mathcal{D}$  is a symmetric tensor of rank 2. Each symmetric second rank tensor can be reduced to diagonal form.

## 4.2. Unnumbered, Numbered, and Bullet Lists

Theoretical models which permit the calculation of the composition dependent from the deeper principles using, statistical mechanics are nowadays still not broadly available. Then the strategy illustrated in the previous section of calculating the concentration for certain initial and boundary conditions is not applicable.

1. This is Numbered list. The atomic mechanisms of diffusion in crystalline materials are closely connected with defects.
2. This is Numbered list. The atomic mechanisms of diffusion in crystalline materials are closely connected with defects.
3. This is Numbered list. The atomic mechanisms of diffusion in crystalline materials are closely connected with defects.

Theoretical models which permit the calculation of the composition dependent from the deeper principles using, statistical mechanics are nowadays still not broadly available. Then the strategy illustrated in the previous section of calculating the concentration for certain initial and boundary conditions is not applicable.

1. This is Unnumbered list. The atomic mechanisms of diffusion in crystalline materials are closely connected with defects.
2. This is Unnumbered list. The atomic mechanisms of diffusion in crystalline materials are closely connected with defects.
3. This is Unnumbered list. The atomic mechanisms of diffusion in crystalline materials are closely connected with defects.

This chapter will concentrate on bulk diffusion in solid metals and alloys. Most of the solid elements are metals. Furthermore, diffusion properties and atomic mechanisms of diffusion have most thoroughly been investigated in metallic solids.

- This is Bullet list. The atomic mechanisms of diffusion in crystalline materials are closely connected with defects.
- This is Bullet list. The atomic mechanisms of diffusion in crystalline materials are closely connected with defects.
- This is Bullet list. The atomic mechanisms of diffusion in crystalline materials are closely connected with defects.

This chapter will concentrate on bulk diffusion in solid metals and alloys. Most of the solid elements are metals. Furthermore, diffusion properties and atomic mechanisms of diffusion have most thoroughly been investigated in metallic solids.

- 1 This is Nested listing - 1
  - This is Nested listing - A
  - This is Nested listing - B
- 5 This is Nested listing - 2
  - This is Nested listing - A
  - This is Nested listing - B
1. This is Nested listing - 3

### 4.3. Descriptions and quotes

Typical examples are nucleation of new phases, diffusive phase transformations, precipitation and dissolution of a second phase, recrystallization, high-temperature creep, and thermal oxidation.

**First entry** description text description text description text description text description text description text

**Second long entry** description text description text description text description text description text description text description text

**Third entry** description text description text description text description text description text

**Fourth entry** description text description text

This chapter will concentrate on bulk diffusion in solid metals and alloys. Most of the solid elements are metals. Furthermore, diffusion properties and atomic mechanisms of diffusion have most thoroughly been investigated in metallic solids.

**Quote:** The atomic mechanisms of diffusion in crystalline materials are closely connected with diffusion in crystalline materials are closely connected with defects.

This chapter will concentrate on bulk diffusion in solid metals and alloys. Most of the solid elements are metals. Furthermore, diffusion properties and atomic mechanisms of diffusion have most thoroughly been investigated in metallic solids.

### 4.4. Numbered Display Equations

**WileySTAT-V1** class file makes the full functionality of  $\text{\LaTeX}$  available. We encourage the use of the `align`, `gather` and `multiline` environments for displayed mathematics.

This chapter will concentrate on bulk diffusion in solid metals and alloys. Furthermore, diffusion properties and atomic mechanisms of diffusion have most thoroughly been investigated in metallic solids.

$$-\frac{D_1}{V} \left. \frac{\partial C}{\partial x} \right|_{x=0} = \begin{cases} C_1(y) - \eta_\alpha C_\alpha & \text{for } \alpha \text{ phase} \\ C_2(y) - \eta_\alpha C_\alpha & \text{for } \beta \text{ phase} \end{cases} \quad (1)$$

Equation (1) and Eqs. 1–4 implies that  $\mathcal{D}$  varies with direction. In general the diffusion flux and the concentration gradient are not always antiparallel. They are antiparallel for an isotropic medium.

$$C(x, y) = B_1 \exp \left( -\frac{V}{D} x \right) + \sum_{n=1}^{\infty} B_n \cos \frac{n\pi y}{L} \quad (2)$$

$$C(x, y) = B_2 \exp \left\{ \left[ -\frac{V}{2D} - \left[ \left( \frac{V}{2D} \right)^2 + \left( \frac{n\pi}{L} \right)^2 \right]^{1/2} \right] x \right\} \quad (3)$$

For anisotropic media and non-cubic crystalline solids  $\mathcal{D}$  is a symmetric tensor of rank 2 Benson (1992). Each symmetric second rank tensor can be reduced to diagonal form. The diffusion flux is antiparallel to the concentration

gradient only for diffusion along the orthogonal principal directions. If  $x_1, x_2, x_3$  denote these directions and  $j_1, j_2, j_3$  the pertaining components of the diffusion flux, Eq. (2) can be written as

$$\begin{aligned} & [\mathcal{D}d^{sn}D_n(\nabla \times \tilde{\mathbf{u}}^s) + D_n(\nabla \times \mathbf{u}^s) -] \delta \mathbf{u}^s(\mathbf{x}, t) \\ & = \lim_{\alpha \rightarrow 0} \{ D_n(\nabla \times \tilde{\mathbf{u}}^s) + D_n(\nabla \times \mathbf{u}^s) - [\nabla \times (\mathbf{u}^s + \alpha \delta \mathbf{u}^s)] - D_n(\nabla \times \tilde{\mathbf{u}}^s) [\nabla \times \mathbf{u}^s] \} / \alpha. \end{aligned} \quad (4)$$

where diffusion coefficient for a direction  $D_I, D_{II}, D_{III}$  denote the three principal diffusivities. The diffusion coefficient for a direction  $(\alpha_1, \alpha_2, \alpha_3)$  is obtained from

$$\begin{aligned} s(nT_s) &= s(t) \times \sum_{n=0}^{N-1} \delta(t - nT_s) \leftrightarrow \text{DFTS} \left( \frac{m}{NT_s} \right) \\ &= \frac{1}{N} \sum_{n=0}^{N-1} \sum_{k=-N/2}^{N/2-1} s_k e^{j2\pi k \Delta f n T_s} e^{-j\frac{2\pi}{N} mn} \end{aligned} \quad (5)$$

#### 4.5. Unnumbered Inline and Display Equations

Although there will be a gradient in the concentration of the trace element, its total concentration can be kept so small that the overall composition of the sample during the investigation does practically not change<sup>2</sup>

$$\frac{\partial c}{\partial t} = \frac{\partial}{\partial x} \left( \tilde{D}(c) \frac{\partial c}{\partial x} \right) = \tilde{D}(c) \frac{\partial^2 c}{\partial x^2} + \frac{d\tilde{D}(c)}{dc} \left( \frac{\partial c}{\partial x} \right)^2.$$

The connection between the macroscopically defined tracer self-diffusion coefficient and the atomistic picture of diffusion is the famous Einstein-Smoluchowski relation discussed in detail in Sect. 4.6.

Where  $l$  denotes the jump length and  $\tau$  the mean residence time of an atom on a certain site of the crystal<sup>3</sup>. The quantity  $f$  is the correlation factor. For self-diffusion in cubic crystals  $f$  is a numeric factor.

$$\begin{aligned} U &= T\delta S - P\delta V + \sum_{i=1}^c \mu_i \delta N_i, \\ N &= \sum_{i=1}^c \mu_i \delta N_i + \sum_{i=1}^c \mu_i. \end{aligned}$$

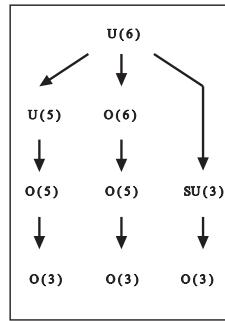
where diffusion coefficient for a direction  $D_I, D_{II}, D_{III}$  denote the three principal diffusivities. The diffusion coefficient for a direction  $(\alpha_1, \alpha_2, \alpha_3)$  is obtained from

$$\begin{aligned} c_1 &= A_1 + c_2 A_2 c_3 A_3 \\ c_1 &= A_1 + c_2 A_2 c_3 A_3 \quad \text{with multiple lines without number} \\ c_1 &= A_1 + c_2 A_2 c_3 A_3 \quad \text{with multiple lines without number} \end{aligned}$$

are obtained. These diffusion coefficients are denoted as impurity diffusion coefficients or sometimes also as foreign atom diffusion coefficients.

<sup>2</sup>From an atomistic viewpoint this implies that a tracer atom is not influenced by other tracer atoms.

<sup>3</sup>Equation (4) considers only the simplest case: cubic structure, all sites are energetically equivalent, only jumps to nearest neighbours are allowed.



**Figure 1.** A short description of the figure content should go here. Hypothetical data on the effect of Drug A versus Placebo on the number of study participants with infection present at 1 week, among participants with Infection X.  
Source: This graph was drawn using the Cochrane Collaboration software, Revman (version 5.3).

## 4.6. Figures and Tables

**WileySTAT-V1** class file uses the **graphicx** package for handling figures.

So far we have considered in this section cases where the concentration gradient is the only cause for the flow of matter. We have seen that such situations can be studied using tiny amounts of trace elements in an otherwise homogeneous material. However, from a general viewpoint a diffusion flux is proportional to the gradient of the chemical potential.

For further details on how to size figures, etc., with the **graphicx** package see, for example, Hirt et al. (1974) or (Dukowicz, 1984). If figures are available in an acceptable format (for example, .eps, .ps) they will be used but a printed version should always be provided.

**Table 1.** This is sample table caption. A researcher must decide which observed variables to include in the theoretical model and how these observed variables measure.

Spanned heading 1			Spanned heading 2	
col1 head	col2 head	col3 head	col4 head	col5 head
col1 text	col2 text	col3 text	12.34	col5 text 1
col1 text	col2 text	col3 text	1.62	col5 text 2
col1 text	col2 text	col3 text	51.809	col5 text 3

Unnumbered table footnotes.

**Table 2.** This is sample table caption.

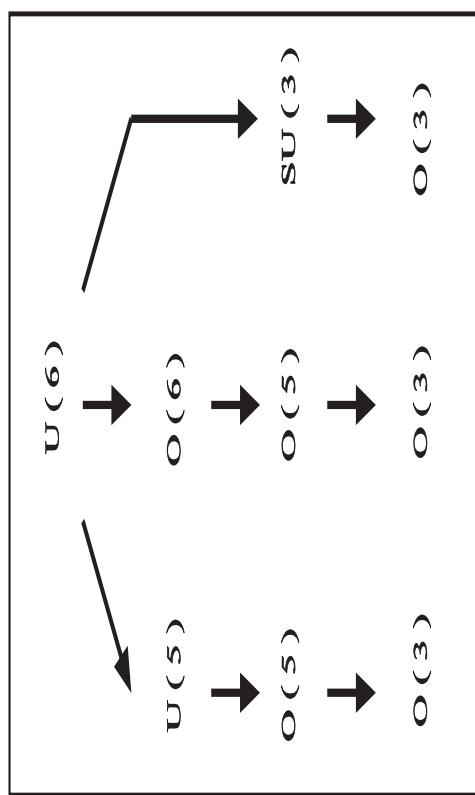
col1 head	col2 head	col3 head	col4 head	col5 head
col1 text	col2 text	col3 text	col4 text	col5 text <sup>†</sup>
col1 text	col2 text	col3 text	col4 text	col5 text
col1 text	col2 text	col3 text	col4 text	col5 text <sup>‡</sup>

<sup>†</sup>Example for a first table footnote.

<sup>‡</sup>Example for a second table footnote.

**Table 3.** Sideways table caption. For decimal alignment refer column 4 to 9 in tabular preamble.

S No	col2 head	col3 head	10	20	30	10	20	30
1	col2 text	col3 text	0.7568	1.0530	1.2642	0.9919	1.3541	1.6108
2	col2 text	col2 text	12.5701	19.6603	25.6809	18.0689	28.4865	37.3011
3	col2 text	col3 text	0.7426	1.0393	1.2507	0.9095	1.2524	1.4958
4	col2 text	col3 text	12.8008	19.9620	26.0324	16.6347	26.0843	34.0765
5	col2 text	col3 text	0.7285	1.0257	1.2374	0.8195	1.1407	1.3691
6	col2 text	col3 text	13.0360	20.2690	26.3895	15.0812	23.4932	30.6060



**Figure 2.** Sideways figure caption. Sideways figure caption.

## 4.7. Examples for Enunciations

This section cases where the concentration gradient is the only cause for the flow of matter. We have seen that such situations can be studied using tiny amounts of trace elements in an otherwise homogeneous material.

**Theorem 1** (Theorem subhead). *We have seen that such situations can be studied using tiny amounts of trace elements in an otherwise homogeneous material. However, from a general viewpoint a diffusion flux is proportional to the gradient of the chemical potential.*

**Proposition 2.** *We have seen that such situations can be studied using tiny amounts of trace elements in an otherwise homogeneous material. However, from a general viewpoint a diffusion flux is proportional to the gradient of the chemical potential.*

However, from a general viewpoint a diffusion flux is proportional to the gradient of the chemical potential, from a general viewpoint a diffusion flux is proportional to the gradient of the chemical potential.

**Proof.** *Example for proof text. Example for proof text.* □

**Definition 3** (Definition sub head). *We have seen that such situations can be studied using tiny amounts of trace elements in an otherwise homogeneous material. However, from a general viewpoint a diffusion flux is proportional to the gradient of the chemical potential.*

**Example 4.** *We have seen that such situations can be studied using tiny amounts of trace elements in an otherwise homogeneous material. However, from a general viewpoint a diffusion flux is proportional to the gradient of the chemical potential.*

**Proof of Theorem 1.** *In  $G$  denotes Gibbs free energy,  $n_i$  the number of moles of species  $i$ ,  $T$  the temperature, and  $p$  the hydrostatic pressure. The chemical potential depends on the alloy composition.* □

## 4.8. Code Display

Use `\begin{verbatim}... \end{verbatim}` for program codes without math. Use `\begin{alltt}... \end{alltt}` for program codes with math. Based on the text provided inside the optional argument of `\begin{code}[Psecode|Listing|Box|Code|Specification|Procedure|Sourcecode|Program]... \end{code}` tag corresponding boxed like floats are generated. Also note that `\begin{code}[Code|Listing]... \end{code}` tag with either Code or Listing text as optional argument text are set with computer modern typewriter font. All other code environments are set with normal text font. Refer below example:

Listing 1: Descriptive Caption Text

```
for i:=maxint to 0 do
begin
{ do nothing }
end;
Write('Case insensitive ');
Write('Pascal keywords.');
```

## 4.9. Algorithms

Below is Algorithm example. Use algorithm package documentation for more details:

---

**Algorithm 1** Pseudocode for our algorithm

---

```
for e doach frame
    for w doater particles  $f_i$ 
        compute fluid flow
        compute fluid–solid interaction
        apply adhesion and surface tension
    end for
end for
```

---

## 4.10. Cross-referencing

The use of the L<sup>A</sup>T<sub>E</sub>X cross-reference system for figures, tables, equations, etc., is encouraged (using `\ref{<name>}` and `\label{<name>}`).

## 4.11. Acknowledgements

An Acknowledgements section is started with `\ack` or `\acks` for *Acknowledgement* or *Acknowledgements*, respectively. It must be placed just before the References.

## 4.12. Bibliography

The bibliography section is using the standard "natbib" package for author-year citation. The normal commands for producing the reference list are:

where `\bibitem{x-ref label}` corresponds to `\citet{x-ref label}` (direct citation) or `\citep{x-ref label}` (indirect citation) in the body of the article.

For those authors that are using BIBT<sub>E</sub>X, **wb\_stat.bst** is included in the zip package as well. This bibliography style will format the reference based on *Stat* requirement (as seen in the reference section).

## 4.13. Visuanimation

Visuanimation is a term coined for visualization through animations, please refer CTAN for L<sup>A</sup>T<sub>E</sub>X packages. It requires the package *animate* and *multimedia*. It allows to embed animations in the paper itself and to store larger movies in the online supplemental material. Examples of statistics research projects using a variety of visuanimations range from exploratory data analysis of image data sets to spatio-temporal extreme event modelling, from multiscale analysis of classification methods to the study of the effects of a simulated explosive volcanic eruption and emulation of climate model output, from spatio-temporal wind roses to point processes on the sphere. The use of visuanimations in *Stat* papers is highly encouraged.

## 4.14. Copyright Statement

Please be aware that the use of this  $\text{\LaTeX} 2\epsilon$  class file is governed by the following conditions. Copyright © 2018 John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK. All rights reserved.

## 4.15. Rules of Use

This class file is made available for use by authors who wish to prepare an article for publication in *Stat* published by John Wiley & Sons, Ltd. The user may not exploit any part of the class file commercially.

This class file is provided on an *as is* basis, without warranties of any kind, either express or implied, including but not limited to warranties of title, or implied warranties of merchantability or fitness for a particular purpose. There will be no duty on the author[s] of the software or John Wiley & Sons, Ltd to correct any errors or defects in the software. Any statutory rights you may have remain unaffected by your acceptance of these rules of use.

## 4.16. The Three Golden Rules

Before we proceed, we would like to stress *three golden rules* that need to be followed to enable the most efficient use of your code at the typesetting stage:

- (i) keep your own macros to an absolute minimum;
- (ii) as  $\text{\TeX}$  is designed to make sensible spacing decisions by itself, do *not* use explicit horizontal or vertical spacing commands, except in a few accepted (mostly mathematical) situations, such as  $\backslash$ , before a differential  $d$ , or  $\backslashquad$  to separate an equation from its qualifier;
- (iii) follow this sample for styles

## Acknowledgement

I would like to thank....

## A. Appendix Examples

A researcher must decide which observed variables to include in the theoretical model and how these observed variables measure latent variables.

$$\|\tilde{X}(k)\|^2 = \frac{\left\| \sum_{i=1}^p \tilde{Y}_i(k) + \sum_{j=1}^q \tilde{Z}_j(k) \right\|^2}{(p+q)^2} \quad (\text{A1})$$

$$\|\tilde{Y}(k)\|^2 = \frac{\sum_{i=1}^p \|\tilde{Y}_i(k)\|^2 + \sum_{j=1}^q \|\tilde{Z}_j(k)\|^2}{p+q}. \quad (\text{A2})$$

**Table A1.** This is sample appendix table caption.

col1 head	col2 head	col3 head	col4 head	col5 head
col1 text	col2 text	col3 text	12.34	col5 text 1
col1 text	col2 text	col3 text	1.62	col5 text 2
col1 text	col2 text	col3 text	51.809	col5 text 3

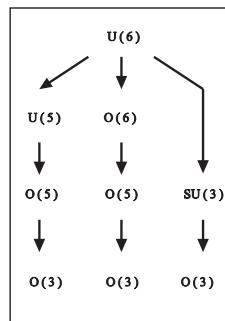
Unnumbered table footnotes.

## B. Theoretical Model

A researcher must decide which observed variables to include in the theoretical model and how these observed variables measure latent variables.

$$D(\alpha_1, \alpha_2, \alpha_3) = \alpha_1^2 D_I + \alpha_2^2 D_{II} + \alpha_3^2 D_{III}, \quad (\text{B1})$$

Although there will be a gradient in the concentration of the trace element, its total concentration can be kept so small that the overall composition of the sample during the investigation does practically not change.

**Figure B1.** A short description of appendix figure content should go here.

## Related Articles

**Scoring Rules; Subjective Probabilities: Theory; Subjective Randomness; Subjective Probability and Human Judgement; Subjective Probabilities: Overview; Subjective Expected Utility; Expert Opinion in Reliability; Expert Judgment; Expert Elicitation for Risk Assessment; Prior Distribution Elicitation.**

## Further Reading

1. Knuth, D. E. (1986). *The TeXbook*. Reading, MA: Addison-Wesley.
2. Lamport L. 1994. LaTeX is document preparation system, commonly used by scientists engineers, mathematicians, and professionals. *L<sup>A</sup>T<sub>E</sub>X: A Document Preparation System* (2nd edn). Addison-Wesley.
3. Mittelbach F, Goossens M. 2004. *The L<sup>A</sup>T<sub>E</sub>X Companion* (2nd edn). Addison-Wesley.