

# LatFit - Producing high accuracy lattice models from protein atomic co-ordinates including side chains

Martin Mann<sup>\*,1</sup>, Rhodri Saunders<sup>2</sup>, Cameron Smith<sup>1</sup>, Rolf Backofen<sup>1</sup> and Charlotte M. Deane<sup>2</sup>

<sup>1</sup>University of Freiburg, Bioinformatics Group, Georges-Köhler-Allee 106, 79110 Freiburg, Germany

<sup>2</sup>University of Oxford, Department of Statistics, 1 South Parks Road, Oxford, England, OX1 3TG

Email: MM\* - [mmann@informatik.uni-freiburg.de](mailto:mmann@informatik.uni-freiburg.de); RS - [saunders@stats.ox.ac.uk](mailto:saunders@stats.ox.ac.uk); CS - [smithc@informatik.uni-freiburg.de](mailto:smithc@informatik.uni-freiburg.de); RB - [backofen@informatik.uni-freiburg.de](mailto:backofen@informatik.uni-freiburg.de); CD - [deane@stats.ox.ac.uk](mailto:deane@stats.ox.ac.uk);

\*Corresponding author

## Abstract

---

**Background:** Lattice models are a common abstraction used in the study of protein structures and related topics. Various approaches to the *protein chain lattice fitting* problem have been suggested but only one tool is currently openly available and this is for backbone-only models.

**Results:** We introduce LatFit, a new tool to produce high accuracy lattice protein models from protein atomic co-ordinates. It generates backbone-only models as well as lattice models including side chains in various lattice types. LatFit implements a new dRMSD-optimisation fitting procedure in addition to a known cRMSD-optimising method. The program is freely available for download or as a web server at <http://cpsp.informatik.uni-freiburg.de>

**Conclusions:** We model a large non-redundant set of high resolution proteins (SCOP database) on three commonly used lattices: 3D cubic (100), face centred cubic (FCC, 110), and knight's walk (210). Both backbone and side chain models produced by LatFit show low deviation from the original data (e.g. 1.5Å RMSD in FCC lattice). To our knowledge, this is the first comprehensive study of lattice quality for on-lattice protein models including side chains.

---

## Background

It is not always computationally feasible to undertake protein structure studies using full atom representations. The challenge is to reduce complexity while maintaining detail [1–3]. Lattice protein models are often used to achieve this but in general only the protein backbone or the amino acid centre of mass is represented [4–11]. A huge variety of lattices and energy functions have previously been

developed [4, 12, 13].

In order to evaluate the applicability of different lattices and to enable the transformation of real protein structures into lattice models, a representative lattice protein structure has to be calculated. Mañuch and Gaur have shown the NP-completeness of this problem for backbone-only models in the 3D-cubic lattice and named it the *protein chain lattice fitting (PCLF) problem* [14].

The PCLF problem has been widely studied for

backbone-only models [12, 15–22]. The most important aspects in producing lattice protein models with a low root mean squared deviation (RMSD) are the lattice co-ordination number and the neighbourhood vector angles [17, 22]. Lattices with intermediate co-ordination numbers, such as the face-centred cubic (FCC) lattice, can produce high resolution backbone models [17] and have been used in many protein structure studies (e.g. [3, 23, 24]). However, the use of backbone models is limited since they do not account for the space required for side chain packing.

To overcome this restriction lattice protein models that include side chains have been introduced [25–31]. Reva *et al.* have to our knowledge developed the only approach to solve the PCLF problem [30]. The CABS-tools by Kolinski and co-workers utilize a hybrid on-lattice (backbone) and off-lattice (side chain) protein representation does not attempt to answer the PCLF problem [29, 32].

In this manuscript we use the side chain model definition of Bromberg and Dill [26], where each amino acid is represented by two on-lattice monomers: one represents the side chain atoms and one the  $C_\alpha$  atom. This explicit representation of side chains prevents unnatural collapse during structural studies [33] and enables the reconstruction of full atom protein data [34].

To the best of our knowledge, there is only one publicly available approach, namely **LocalMove**, to derive lattice protein models from real proteins despite a large number of published methods. **LocalMove** was introduced by Ponty *et al.* for backbone-only models in 3D-cubic and FCC lattice [21].

We present our tool **LatFit** to tackle this lack of available implementations. **LatFit** solves the PCLF problem, i.e. transforms a protein from full atom co-ordinate data to a lattice model, and is available as both a stand-alone tool for high-throughput pipelines and a web interface for *ad hoc* usage. A new fitting procedure that optimises distance RMSD enables rotation independent lattice model creation of protein structures. The method is applicable to arbitrary lattices and handles both backbone and side chain representations with equivalent accuracy. A depiction of the workflow is given in Fig. 1.

Utilising **LatFit** we present the first comprehensive study of lattice quality for protein models including side chains. In our test, **LatFit** fitted the majority of models on an FCC lattice within 1.5Å RMSD.

## Methods

In order to enable a precise formulation of the method we introduce some preliminary definitions. A lattice  $L$  is defined by a set of neighboring vectors  $\vec{n} \in N$  of equal length ( $\forall_{\vec{n}_i, \vec{n}_j \in N} : |\vec{n}_i| = |\vec{n}_j|$ ), each with a reverse ( $\forall_{\vec{n} \in N} : -\vec{n} \in N$ ), such that  $L = \{\vec{x} \mid \vec{x} = \sum_{\vec{n}_i \in N} d_i \cdot \vec{n}_i \wedge d_i \in \mathbb{Z}_0^+\}$ .  $|N|$  gives the coordinate number of the lattice, e.g. 6 for 3D-cubic or 12 for the FCC lattice. A lattice protein structure with side chains of length  $l$  is defined by a sequence of lattice nodes  $M^b = (M_1^b, \dots, M_l^b) \in L^l$  representing the backbone monomers of each amino acid and the according sequence  $M^s = (M_1^s, \dots, M_l^s) \in L^l$  for their side chain position. A valid structure ensures backbone connectivity ( $\forall_{i < l} : M_i^b - M_{i+1}^b \in N$ ), side chain connectivity ( $\forall_i : M_i^b - M_i^s \in N$ ), as well as selfavoidance ( $\forall_{i \neq j} : M_i^b \neq M_j^b \wedge M_i^s \neq M_j^s$  and  $\forall_{i,j} : M_i^b \neq M_j^s$ ).

## Fitting Procedure

Given a protein structure of length  $l$  in Protein Database (PDB) format [35], **LatFit** builds up the lattice protein sequentially, one amino acid at a time, starting from the amino-terminus.

First, all neighboring vectors  $\vec{n} \in N$  of the used lattice  $L$  are scaled to a length of 3.8Å, which is the mean distance between consecutive  $C_\alpha$  atoms and close to the mean distance between a  $C_\alpha$  atom and the associated side chain centroid ( $\approx 3.6\text{Å}$ ). This scaling enables a reasonable mapping of the protein into the lattice; since each amino acid will be represented by two monomers while connected monomers will have distance  $|\vec{n}| = 3.8\text{Å}$ . Therefore, all resulting measures will be directly interpretable in Å units.

The positions for each amino acid  $i$  to be fitted, i.e. the  $C_\alpha$  position of the backbone  $P_i^b$ , and the centroid  $P_i^s$  (geometric center) of all non-hydrogen atom co-ordinates of the side chain, are extracted from the PDB file.

The lattice model is derived by one of the following procedures optimising either a distance or coordinate RMSD. Both methods are introduced for lattice proteins including side chains but can be used to derive backbone-only lattice models as well. A sketch of the fitting workflow is given in Fig. 1.

### dRMSD Optimisation

The fitting follows a greedy iterative chain-growth procedure. The initial lattice model’s backbone and side chain position ( $M_1^b$  and  $M_1^s$ ) are placed arbitrarily but adjacent ( $M_1^b - M_1^s \in N$ ). For each iteration  $1 < i \leq l$ , all valid placements of the next  $M_i^b$  and  $M_i^s$  on the lattice are calculated. A distance RMSD (dRMSD, Eqn. 1) evaluation is used to identify the best  $n_{keep}$  structures of length  $i$  for the next extension iteration. Since dRMSD is a rotation/reflection independent measure, symmetric structures must be filtered.

To calculate the final fit of the initial protein  $P$ , a superpositioning of the dRMSD-optimised structure  $M$  and a reflected version  $M'$  is done using the method by Kabsch [36]. The superpositioning translates and rotates  $M/M'$  in order to achieve the best mapping onto  $P$ . The superpositioning with lowest co-ordinate RMSD (cRMSD, Eqn. 2) is selected and finally returned.

$$\text{dRMSD} = \sqrt{\frac{\sum_{i < j} (|P_i - P_j| - |M_i - M_j|)^2}{l \cdot ((2 \cdot l) - 1)}} \quad (1)$$

with  $P = P^s \cup P^b$ , and  $M = M^s \cup M^b$ .

$$\text{cRMSD} = \sqrt{\frac{\sum_{i=1}^l (|P_i^b - M_i^b|)^2 + (|P_i^s - M_i^s|)^2}{2 \cdot l}} \quad (2)$$

### cRMSD Optimisation

A cRMSD evaluation according to Eq. 2 depends on the superpositioning of the protein and its model. Thus the best relative lattice orientation has to be identified in addition to the best model. Once the orientation is fixed, a cRMSD evaluation allows for a fast, additive RMSD update along the chain extension.

We implement a cRMSD optimising method following [6, 17] as an alternative fitting strategy. In general a user defined number of rotation intervals  $r$  are performed for each of the XYZ rotation axes. For each rotation, we transform  $P^b$  and  $P^s$  into  $\hat{P}^b$  and  $\hat{P}^s$ , respectively, to obtain the rotated current target structure.

The fitting procedure follows a chain-growth approach:  $P_1^b$  is placed onto an arbitrary lattice node  $M_1^b$  and  $P_1^s$  to the closest representative node  $M_1^s$  adjacent to  $M_1^b$ . Now, all valid placements of the next  $M_i^b$  and  $M_i^s$  on the lattice are calculated. Using the co-ordinate RMSD (cRMSD, Eqn. 2) we evaluate all

derived models and keep the best  $n_{keep}$  for the next extension following [17] until all amino acids have been placed.

By applying the above cRMSD based fitting procedure we obtain the best fit for the current rotation. An iterative application of this procedure then results in the overall best fit for all screened rotations. Since our screen of XYZ rotations was discretised, the current rotation might be refineable. Therefore, another rotational refinement can be applied that investigates  $r^{ref}$  small rotation intervals around the best rotation from the first screen [6].

The run time of the cRMSD-method scales with respect to the lattice co-ordination number,  $n_{keep}$ , and most importantly the number of rotation intervals  $r$  and  $r^{ref}$  considered.

### Futher Features

Coordinate data in the PDB is often incomplete. For example flexible loop structures are hard to resolve by current methods [37]. This results in missing co-ordinate data for certain substructures within PDB files. **LatFit** enables a structural fitting of even such fragmented PDB structures. It produces a lattice protein fragment for each fragment of the original protein while ensuring that all are placed in the same lattice orientation.

Currently, **LatFit** supports the 2D-square, 3D-cubic (100), 3D face centered cubic (FCC, 110) and 3D knights walk (210) lattice. The modular software design of our open source program enables an easy and straight forward implementation of other lattices.

Supported output formats of **LatFit** are the PDB format, the Chemical Markup Language (CML) format, as well as a simple XYZ coordinate output. A highly compact string representation of the lattice protein is also given in absolute move strings that encode the series of neighboring vectors  $\vec{n} \in N$  along the structure.

The generated absolute move string can be directly used to apply other lattice protein tools onto the resulting structures, e.g. from the CPSP-package for HP-type lattice protein models [10, 38] or from the **LatPack** tools for arbitrary lattice models [11, 39].

The web interface of **LatFit**, integrated into the CPSP-web-tools [38], enables *ad hoc* usage of the tool. Either a protein structure in PDB format can be uploaded or a valid identifier from the PDB database

given.

Our default parameters enable a direct application of **LatFit** resulting in a balanced tradeoff between runtime and fitting quality. The computations are done remotely on a computation cluster while the user can trace the processing status via the provided job identifier and according link. Results are available and stored for 30 days.

Results can be visualised using **Jmol** [40] for an interactive presentation of the final protein structure. The output file is also available for download. The final dRMSD and cRMSD values of the lattice protein compared to the original protein are given as well as the absolute move string encoding of the resulting structure. For an example of the **LatFit** web interface see Fig. 2.

Further details regarding the methods implemented, the output formats supported and the applicable parameterisation are located in the **LatFit** manual distributed with the source code. We provide an extensive help page and a frequently asked questions (FAQ) section within the web interface.

<http://www.bioinf.uni-freiburg.de/Software/LatPack/>

A web interface for *ad hoc* usage is available at

<http://cpsp.informatik.uni-freiburg.de>

## Results and Discussion

We use **LatFit** to derive protein models on the commonly used 3D cubic, FCC, and knights walk lattices [17]. Our test set was taken from the PISCES webservice [41] on March 10<sup>th</sup> 2009 (40% sequence identity cut-off, chain length 50-300, R-factor  $\leq 0.3$  and resolution  $\leq 1.5\text{\AA}$ ). Given our requirement for side chains,  $C_\alpha$ -only chains were ignored. The resulting benchmark set contains 1198 proteins exhibiting a mean length of 160 ( $\sigma = 64$ ).

In accordance with previous studies [17], cRMSD and dRMSD are used to assess model quality. cRMSD measures the similarity in co-ordinate position whereas dRMSD measures the similarity of interatomic distances. Due to the scaling of our lattice, RMSD results are in  $\text{\AA}$  rather than the scaled values provided by Ponty et al. [21].

Each protein was fitted twice onto the lattice using either our dRMSD or cRMSD-optimising method. dRMSD-optimisation was parameterised with  $n_{keep} = 1000$ . For cRMSD-optimising runs, we used the parameters  $r = 10$  and  $r^{ref} = 5$  for

backbone-only fits, and  $r = 5$  and  $r^{ref} = 3$  for side chain fits. A rotation range of  $[0, \frac{\pi}{2}]$  and  $n_{keep} = 5$  was used for initial rotations. Rotational refinement was applied onto the interval  $\pm[0, \frac{\pi}{10}]$  around the best initial rotation to derive the final fit.

Our backbone model RMSD values presented in Table 1 are competitive or superior to known fitting results from the literature [6, 12, 17] and reproduce the high quality previously achieved by other methods using the FCC and 210 lattices.

**LatFit** is designed for side chain models and results here are strong (see Table 2). In general, side chain models produce slightly larger RMSD values than the equivalent backbone model. This is due to the fact that the variation in distance between consecutive  $C_\alpha$  atoms (fitted in both models) is lower than that between  $C_\alpha$  atoms and their side chain centroid (fitted only in side chain models). In lattice models every distance is fixed at  $3.8\text{\AA}$  which results in a higher mean displacement of the side chain. Nevertheless, high accuracy fits are still attained. Results in our test set have mean dRMSDs of about 1.5 and  $1.2\text{\AA}$  in the FCC and 210 lattice respectively. The strength of **LatFit** is its ability to produce both side chain and backbone-only lattice protein models. High accuracy models can be produced on the FCC lattice in seconds to minutes. Fits on the 210 lattice take orders of magnitude longer for relatively little gain in model accuracy. For this reason we recommend using the FCC lattice for detailed high-throughput protein structure studies in both backbone-only and side chain representing lattice models.

## Conclusion

**LatFit** enables the automated high resolution fitting of both backbone and side chain lattice protein models from full atomic data in PDB format. We demonstrate its high accuracy on three widely used lattices using a large, non-redundant protein data set of high resolution. Side chain fits show on average a higher deviation than backbone models, but both produce high quality fits with results generally less than  $1.5\text{\AA}$  on the face-centred cubic lattice. To our knowledge, this is the first study and publicly available implementation for side chain models in this field. Available via web interface and as a stand-alone tool, **LatFit** addresses the lack of available programs and is well placed to enable further,

more detailed investigation of protein structure in a reduced complexity environment.

## Availability and Requirements

**Project name:** LatFit

**Project home page:** Web interface and source at <http://cpsp.informatik.uni-freiburg.de>  
<http://www.bioinf.uni-freiburg.de/Software/LatPack/>

**Operating system(s):** Web interface: all common browser supporting Javascript and Java applets (for visualisation); Source: all Linux based systems (including Cygwin for MS Windows<sup>TM</sup>)

**Programming language:** Web interface: Java Server Pages (JSP) and Javascript; Source: C++

**Other requirements:** BIU C++ library is part of the source package.

**License:** BSD-style license

**Any restrictions to use by non-academics:** none

## Authors contributions

Idea and method design by MM, RS, and CD. Data set compiled by RS. Implementation and verification done by MM. CS and MM implemented the web interface. All authors contributed to and have approved the final manuscript.

## References

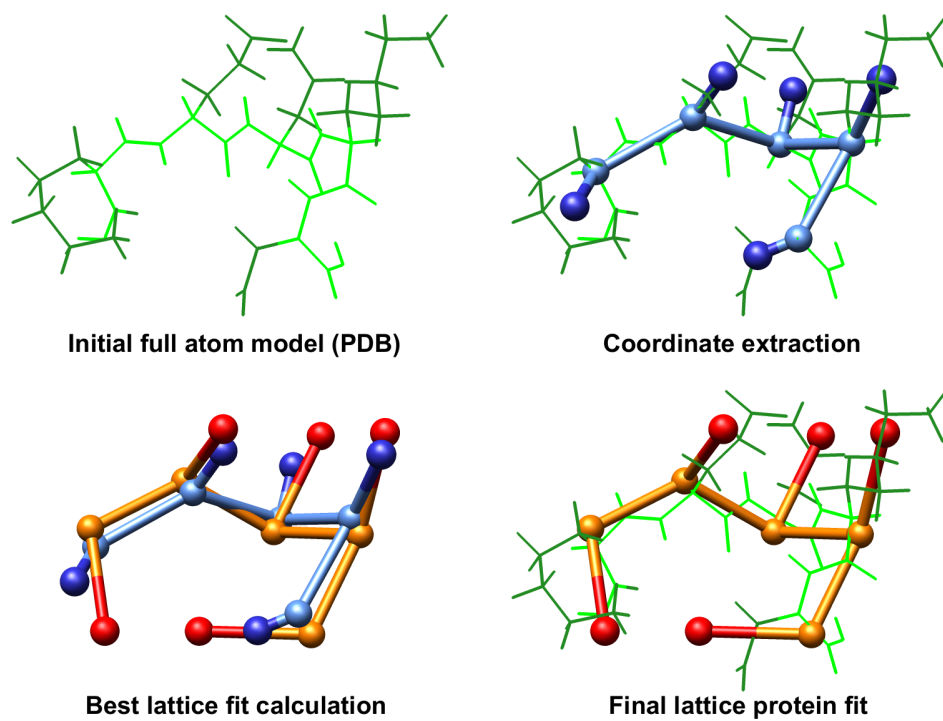
1. Mirny L, Shakhnovich E: **Protein Folding Theory: From Lattice to All-Atom Models.** *Annual Review of Biophysics and Biomolecular Structure* 2001, **30**:361–396.
2. Dill KA, Ozkan SB, Shell MS, Weikl TR: **The Protein Folding Problem.** *Annual Review of Biophysics* 2008, **37**:289–316.
3. Istrail S, Lam F: **Combinatorial Algorithms for Protein Folding in Lattice Models: A Survey of Mathematical Results.** *Commun. Inf. Syst.* 2009, **9**(4):303–346.
4. Dill KA: **Theory for the folding and stability of globular proteins.** *Biochemistry* 1985, **24**(6):1501–9.
5. Renner A, Bornberg-Bauer E: **Exploring the fitness landscapes of lattice proteins.** In *Pac Symp Biocomput.* 1997:361–372.
6. Miao J, Kleinseetharaman J, Meirovitch H: **The Optimal Fraction of Hydrophobic Residues Required to Ensure Protein Collapse.** *J Mol. Bio.* 2004, **344**(3):797–811.
7. Backofen R, Will S: **A Constraint-Based Approach to Fast and Exact Structure Prediction in Three-Dimensional Protein Models.** *Constraints* 2006, **11**:5–30.
8. Huard FPE, Deane CM, Wood GR: **Modelling sequential protein folding under kinetic control.** *Bioinformatics* 2006, **22**(14):e203–210.
9. Deane CM, Dong M, Huard FP, Lance BK, Wood GR: **Cotranslational protein folding—fact or fiction?** *Bioinformatics* 2007, **23**(13):i142–148.
10. Mann M, Will S, Backofen R: **CPSP-tools - Exact and Complete Algorithms for High-throughput 3D Lattice Protein Studies.** *BMC Bioinf* 2008, **9**:230.
11. Mann M, Maticzka D, Saunders R, Backofen R: **Classifying protein-like sequences in arbitrary lattice protein models using LatPack.** *HFSP J* 2008, **2**:396.
12. Godzik A, Kolinski A, Skolnick J: **Lattice representations of globular proteins: How good are they?** *J Comp. Chem.* 1993, **14**(10):1194–1202.
13. Reva BA, Sanner MF, Olson AJ, Finkelstein AV: **Lattice modeling: Accuracy of energy calculations.** *Journal of Computational Chemistry* 1996, **17**(8):1025 – 1032.
14. Mañuch J, Gaur DR: **Fitting protein chains to cubic lattice is NP-complete.** *Journal of bioinformatics and computational biology* 2008, **6**:93–106.
15. Covell DG, Jernigan RL: **Conformations of folded proteins in restricted spaces.** *Biochemistry* 1990, **29**(13):3287–3294.
16. Hinds DA, Levitt M: **A lattice model for protein structure prediction at low resolution.** *Proc Natl Acad Sci USA* 1992, **89**(7):2536–2540.
17. Park BH, Levitt M: **The complexity and accuracy of discrete state models of protein structure.** *J Mol Biol* 1995, **249**:493–507.
18. Rykunov DS, Reva BA, Finkelstein AV: **Accurate general method for lattice approximation of three-dimensional structure of a chain molecule.** *Proteins* 1995, **22**(2):100–109.
19. Reva BA, Rykunov DS, Finkelstein AV, Skolnick J: **Optimization of Protein Structure on Lattices Using a Self-Consistent Field Approach.** *Journal of Computational Biology* 1998, **5**(3):531–538.
20. Koehl P, Delarue M: **Building protein lattice models using self-consistent mean field theory.** *J. Chem. Phys.* 1998, **108**:9540–9549.
21. Ponty Y, Istrate R, Porcelli E, Clote P: **LocalMove: computing on-lattice fits for biopolymers.** *Nucleic Acids Res* 2008, **36**(2):W216–W222.
22. Pierri CL, Grassi AD, Turi A: **Lattices for ab initio protein structure prediction.** *Proteins* 2008, **73**(2):351–361.
23. Jacob E, Unger R: **A tale of two tails: Why are terminal residues of proteins exposed?** *Bioinformatics* 2007, **23**(2):e225–30.
24. Ullah AD, Kapsokalivas L, Mann M, Steinhöfel K: **Protein Folding Simulation by Two-Stage Optimization.** In *Proc. of ISICA'09, Volume 51 of CCIS* 2009:138–145.

25. Sun S: **Reduced representation model of protein structure prediction: statistical potential and genetic algorithms.** *Proteins* 1993, **2**(5):762–785.
26. Bromberg S, Dill KA: **Side-chain entropy and packing in proteins.** *Protein Sci* 1994, **3**(7):997–1009.
27. Hart WE, Istrail S: **Lattice and off-lattice side chain models of protein folding: linear time structure prediction better than 86% of optimal.** *Journal of Computational Biology* 1997, **4**(3):241–59.
28. Heun V: **Approximate protein folding in the HP side chain model on extended cubic lattices.** *Discrete Appl. Math.* 2003, **127**:163–177.
29. Kolinski A, Skolnick J: **Reduced models of proteins and their applications.** *Polymer* 2004, **45**(2):511 – 524.
30. Reva BA, Rykunov DS, Olson AJ, Finkelstein AV: **Constructing Lattice Models of Protein Chains with Side Groups.** *Journal of Computational Biology* 1995, **2**(4):527–535.
31. Zhang Y, Arakaki AK, Skolnick J: **TASSER: an automated method for the prediction of protein tertiary structures in CASP6.** *Proteins* 2005, **61** Suppl 7:91–8.
32. Kolinski A: **Protein modeling and structure prediction with a reduced representation.** *Acta biochimica Polonica* 2004, **51**(2):349–372.
33. Eyrych VA, Standley DM, Friesner RA: **Prediction of protein tertiary structure to low resolution: performance for a large and structurally diverse test set.** *J Mol Biol* 1999, **288**(4):725–742.
34. Feig M, Rotkiewicz P, Kolinski A, Skolnick J, Brooks CL: **Accurate reconstruction of all-atom protein representations from side-chain-based low-resolution models.** *Proteins* 2000, **41**:86–97.
35. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE: **The Protein Data Bank.** *Nucl. Acids Res.* 2000, **28**:235–242.
36. Kabsch W: **A discussion of the solution for the best rotation to relate two sets of vectors.** *Acta Crystallographica* 1978, **A34**:827–828.
37. Choi Y, Deane CM: **FREAD revisited: Accurate loop structure prediction using a database search algorithm.** *Proteins* 2010, **78**(6):1431–1440.
38. Mann M, Smith C, Rabbath M, Edwards M, Will S, Backofen R: **CPSP-web-tools: a server for 3D lattice protein studies.** *Bioinformatics* 2009, **25**(5):676–7.
39. Mann M, Hamra MA, Steinhöfel K, Backofen R: **Constraint-based Local Move Definitions for Lattice Protein Models Including Side Chains.** In *Proceedings of the Fifth Workshop on Constraint Based Methods for Bioinformatics (WCB09)* 2009. [ArXiv:0910.3880].
40. Herráez A: **Biomolecules in the Computer: Jmol to the rescue.** *Biochem. Educ* 2006, **34**(4):255–261.
41. Wang G, Dunbrack RL: **PISCES: recent improvements to a PDB sequence culling server.** *Nucleic Acids Res* 2005, **33**(Web Server issue):W94–8.

## Figures

### Figure 1 - Fitting workflow

The diagram depicts the the fitting process of LatFit for side chain models. Original full atom data is given in green, derived coordinates to fit in blue, and the lattice protein model in orange.



**Figure 2 - LatFit web interface**

A screenshot of the LatFit web interface result visualisation.

## LatPack Tools - LatFit Result

**Menu**

[Home](#)

---

**HPstruct**  
structure pred.

**HPconvert**  
PDB, CML, ...

**HPview**  
3D visualization

**HPdeg**  
degeneracy

**HPnnet**  
neutral network

**HPdesign**  
seq. design

---

**LatFit**  
PDB to lattice

---

**Results**  
direct access

---

[Help](#)

[FAQ](#)

Job ID: 6969628

Job 6969628 SUBMITTED @ 13:14:45 UTC+1 on 2010-03-18  
 Job 6969628 STARTED @ 13:15:02 UTC+1 on 2010-03-18  
 Job 6969628 COMPLETED @ 13:15:05 UTC+1 on 2010-03-18  
<http://csp.informatik.uni-freiburg.de/8080/trunk/LatFitResult.jsp?jobid=6969628> (30 day expiry) [Download Results](#)

**Input Parameters**

<a href="#">PDBFile</a>	<a href="#">InputFile_6969628.pdb</a>
<a href="#">Atom</a>	CoM
<a href="#">Chain Identifier</a>	A
<a href="#">Model Number</a>	1
<a href="#">Lattice Protein Type</a>	Include Side-chains
<a href="#">Lattice Form</a>	FCC
<a href="#">CA-CA bond length</a>	3.8
<a href="#">Optimization Mode</a>	dRMSD
<a href="#">Max. to keep per iteration</a>	100
<a href="#">Output Format</a>	PDB
<a href="#">Output points to fit</a>	yes

**Output**

LatFit has produced a [PDB](#) file available for [download](#). Click [here](#) to view the results.

The following distance measures were obtained:

**cRMSD: 1.84493 Angstroms**  
**dRMSD: 1.41052 Angstroms**

---

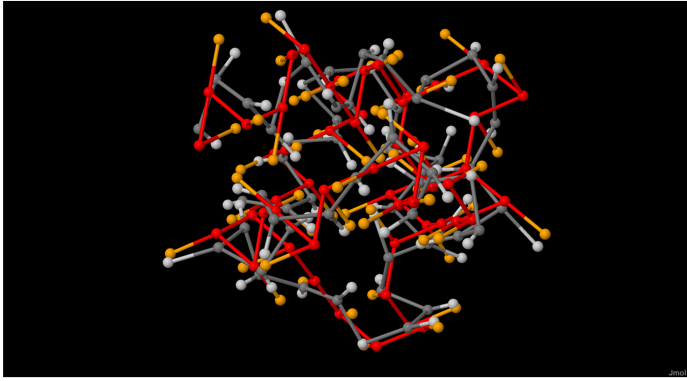
The following absolute move strings have been produced and are available to view:

(BR)BD(BL)FL(LU)FR(BR)BD(BD)LD(BL)FL(RU)FR(BR)BD(BR)LD(FU)LU(BR)BU(FU)LU(LD)BD(RD)RU(RU)B

---

If the molecule does not display, click [here](#) or check the [FAQ](#).

Show LatFit Chain  
 Show Original Chain



■ LatFit Backbone   
 ■ LatFit Sidechain   
 ■ Original Backbone   
 ■ Original Sidechain

---

This result was obtained by using the CPSP-tools package with the following command and arguments:  
 latFit -pdbFile="/scratch/cpsp/bisge000/CPSP\_results/result/2010-03-18\_13~14~45\_latFit\_6969628.pdb"  
 -pdbAtom="CoM" -pdbChain="A" -pdbModel="1" -fitSideChain -lat="FCC" -bondLength="3.8" -opt="D"  
 -nKeep="100" -outMode="PDB" -outOrigData -v

[Legal Disclosure and Contact](#)

## Tables

**Table 1 - Method evaluation for backbone-only models**

The table compares the RMSD mean values ( $\mu$ ) and standard deviations ( $\sigma$ ) from literature to the results from our LatFit cRMSD-optimisation methods for *backbone-only models* on three different lattices.



	Results taken from Park and Levitt [17]		Results taken from Ponty <i>et al.</i> [21]	LatFit cRMSD optimisation	
	dRMSD	cRMSD	cRMSD	dRMSD	cRMSD
	$\mu$	$\mu$	$\mu$ (rescaled to $\text{\AA}$ )	$\mu / \sigma$	$\mu / \sigma$
cub	2.34	2.84	3.5 (0.923 · 3.8)	2.042 / 0.228	2.539 / 0.234
fcc	1.46	1.78	-	1.319 / 0.086	1.641 / 0.090
210	1.02	1.24	-	0.931 / 0.060	1.154 / 0.060

**Table 2 - Method evaluation for side chain models**

The table gives the RMSD mean values ( $\mu$ ) and standard deviations ( $\sigma$ ) of the results from our dRMSD- and the cRMSD-optimisation methods for *side chain models* on three different lattices.

	LatFit - dRMSD optimisation		LatFit - cRMSD optimisation	
	dRMSD	cRMSD	dRMSD	cRMSD
	$\mu / \sigma$	$\mu / \sigma$	$\mu / \sigma$	$\mu / \sigma$
cub	2.779 / 0.754	4.157 / 1.331	2.609 / 0.481	3.286 / 0.624
fcc	1.496 / 0.153	2.104 / 0.246	1.495 / 0.061	1.839 / 0.068
210	1.126 / 0.068	1.601 / 0.100	1.185 / 0.042	1.450 / 0.047