Homology modelling best practices



There are exceptions!



RMSD: ca	1	11	21	31	41
4hhb, chain D	1 VHLTPEEKSA	VTALWGKV.	NVDEVGGEAL	GRLLVVYPWT	QRFFESFG.D
1mbo, chain A RMSD: ca	1 .VLSEGEWQL 51	61	71	81	91
4hhb, chain D	48 L S T P D A V M G N	PKVKAHGKKV	LGA FSDGLAH	LDNLKG.TFA	TLSELHCDKL
1mbo, chain A	49 L K T E A E M K A S	EDLKKHGVTV	LTALGAILKK	KGHH.EAELK	PLAQSHATKH
	101	111	121	131	141
RMSD: ca	101	111	121	131	141
4hhb, chain D	97 HVDPENFRLL	GNVLVCVLAH	HFGKEFTPPV	QAAYQKVVAG	VANALAHKY
1mbo, chain A	98 KIPIKYLEFI	SEAIIHVLHS	RHPGDFGADA	QGAMNKALEL	FRKDIAAKY

Sequence similarity is <25% But identity is high.

Online modelling software (mostly automated)

- <u>https://swissmodel.expasy.org/</u> -We will be using this today.
- <u>https://zhanglab.ccmb.med.umich.edu/I-TASSER/</u> One of the highest ranked for accuracy of produced models.
- <u>http://robetta.bakerlab.org/</u> Online version of ROSETTA protein design suite (still in Beta, long wait times).
- <u>https://salilab.org/modeller/</u> Very popular and a lot of webservers run this.
- <u>https://modbase.compbio.ucsf.edu/modweb/</u> webserver version of modeller.
- <u>http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index</u> works by recognition of conserved folds and also attempts to predict function.

BIOZENTRUM University of Basel The Center for Molecula	SWISS-MODEL	Modelling	Repository	Tools	Documentation		
Start a New M	lodelling Project						
Target Sequence:	Paste your target sequence(s) or UniProtK	B AC here				Supported Inputs	9
(Format must be						Sequence(s)	-
FASTA, Clustal, plain string, or a					1	arget-Template Alignment	-
valid UniProtKB AC)					10	User Template	-
	+ Upload Target Sequence File					DeepView Project	-
Project Title:	Untitled Project						
	Search For Templates	Bui	ld Model				
	By using the SWISS-MODEL server, you agree to comply with the f corresponding articles.	following terms o	of use and to c	cite the			

rget	Target MVVKAVCVINGDAKGTVFFEQESSGTPVKVSGEVCGLAKGLHGFHVHEFGDNTNGCMSSGPHFNPYGKEHGAPVDENRHL 80	Supported Inputs	9
equence(s): Irmat must be	Target GDLGNIEATGDCPTKVNITDSKITLFGADSIIGRTVVVHADADDLGQGGHELSKSTGNAGARIGCGVIGIAKV	Sequence(s)	•
ISTA, Clustal, ain string, or a		Target-Template Alignment	•
lid UniProtKB AC)	Add Hetero Target C Reset	User Template	•
oject Title:	SODC DROME R61851 Superavida diamutada [Cu Zn]	DeepView Project	•
	SODC_DROME Perest Superoxide dismutase [Cu-2n]		
	Search For Templates Build Model		

Start a New M	Nodelling Project @								
Target	Target MVVKAVCVINGDAKGTVFFEQESSGTPVKVSGEVCGLAKGLHGFHVHEFGDNINGCMSSGPHFNPYGKE	HGAPVDENRHL 80	Supported Inputs	9					
Sequence(s): (Format must be	Target GDLGNIEATGDCPTKVNITDSKITLFGADSIIGRTVVVHADADDLGQGGHELSKSTGNAGARIGCGVIG	IAKV 153	Sequence(s)	-					
FASTA, Clustal, plain string, or a			Target-Template Alignment	-					
valid UniProtKB AC)	Add Hetero Target		User Template	•					
			DeepView Project	-					
Project Title:	SODC_DROME P61851 Superoxide dismutase [Cu-Zn]								
	Search For Templates Build Model								
	By using the SWISS-MODEL server, you agree to comply with the following terms of use and to cite the corresponding articles.								
	Note:	f you were to cl	ick build model						
	Swissl	Model would au	Itomatically						
	select	the template							





Summ	nary	Templates 50	Models	li Ł	×					
Temp	olate F	Results o								
Templ	ates	Quaternary St	tructure Sequer	ce Simila	rity	Alignmer	nt of Selecte	ed Templates	5	Build Models 1
More	•									Clear Selection
↓†Sort ₹	Name	≑Title	Coverage	\$GMQE	VQSQE	Identity	Method	¢Oligo State	Ligands	
	3gtv.1.A	Superoxide dismutase [Cu- Zn]		0.80	0.94	60.26	X-ray, 2.2Å	homo-dimer √	2 x ZN 🧉 🛛 👻	6
	3gtv.1.A	Superoxide dismutase [Cu- Zn]		0.82	0.93	61.33	X-ray, 2.2Å	homo-dimer √	2 x ZN ^{II} ♥	
	1n19.1.A	Superoxide Dismutase [Cu-Zn]		0.81	0.92	61.18	X-ray, 1.9Å	homo-dimer √	2 x ZN ^C , ♥ 2 x CU1 c	
	4b3e.1.A	SUPEROXIDE DISMUTASE [CU-ZN]		0.79	0.91	61.18	X-ray, 2.1Å	homo-dimer √	2 x ZN ^{ເਟ} , ♥ 2 x CU ^{ເਟ}	
	4b3e.1.A	SUPEROXIDE DISMUTASE [CU-ZN]		0.81	0.91	61.84	X-ray, 2.1Å	homo-dimer √	2 x ZN [℃] , ♥ 2 x CU [௴]	Z

A side note on template selection

- Your final model is only ever as trustworthy as your template.
- Fortunately SwissModel provides you with a PDB code for the template.
- The PDB has its own set of validation metrics that we can use to asses the model we wish to use.





A Structural View of Biolo

Welcome

주 Deposit

Q Search

Visualize

Analyze

This resource is powered by the Protein Data archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

The RCSB PDB builds upon the data by creating tools and resources for research and education in molecular biology, structural biology, computational biology, and beyond.

New Video: What is a Protein?



Structure Summary	3D View	Annotations	Sequence	Sequence Similarity	Structure Similarity	Experiment	
Biological As	esembly 1	► 1 Str DO Cla Org Exp Mu Dep Dep	N19 ructure of the 1: 10.2210/pdb1 ssification: OX janism(s): Hom pression System tation(s): 3 posited: 2002-1 position Author	e HSOD A4V mutar N19/pdb IDOREDUCTASE o sapiens n: Escherichia coli 0-16 Released: 2002-11- (s): Cardoso, R.M.F., Tha	∎ Displa nt -27 ayer, M.M., DiDonato, M	y Files → ④ Down	load Files -
© 3D View: Structure Ligand Interaction Standalone Viewers Protein Workshop Lig Global Symmetry: Cy Global Stoichiometry	gand Explorer yclic - C2 (3D y: Homo 2-me	nsity Exp Me Res R-V R-V View) rr - A2	berimental Data thod: X-RAY DIf solution: 1.86 Å /alue Free: 0.26 /alue Work: 0.20	a Snapshot w FFRACTION 0 07	WPDB Validation Metric Rfree Clashscore Clashscore Sidechain outliers Sidechain outliers RSRZ outliers Worse Percentile relative Decontile relative	© 3D Report Percentile Ranks to al X-ray structures to X-ray structures of similar resolution	Full Report 0.209 10 0.3% 3.0% 3.9% Better

Structure Summary	3D View	Annotations	Sequence	Sequence Similarity	Structure Similarity	Experiment	_
Biological As	sembly 1	> 1	N19		Display	Files - Ownload Files -	
+		Str DO Cla Org Exp Mut Dep Dep Get	La construction of the suffication: OX panism(s): Homoression System tation(s): 3 posited: 2002-1 position Author 2007, E.D., Taine	e HSOD A4V mutar N19/pdb IDOREDUCTASE to sapiens m: Escherichia coli 0-16 Released: 2002-11 r(s): Cardoso, R.M.F., Th er, J.A.	-27 ayer, M.M., DiDonato, M	., <u>Lo, T.P., Bruns, C.K.,</u>	
© 3D View: Structure Ligand Interaction	Electron Der	nsity Exp	erimental Data	a Snapshot v	wwPDB Validation	O 3D Report Full Report Percentile Ranks Value 0,200	t Ie
Standalone Viewers Protein Workshop Lig	gand Explorer	R-V R-V	alue Free: 0.26 alue Work: 0.26	0 07	Clashscore Clashscore	10 0.3% 3.0%	6
Global Symmetry: Cy Global Stoichiometry	vclic - C2 (3D v: Homo 2-me	View) r - A2			RSRZ outliers	Better a # X-ray structures a X-ray structures of similar resolution	5

Structure Summary	3D View	Annotations	Sequence	Sequence Similarity	Structure Similarity	Experiment
Biological As	sembly 1	> 1	N19		Display	/ Files -
÷		Str DOI Cla Org Exp Mut Dep Get	ructure of th I: 10.2210/pdb1 ssification: <u>OX</u> panism(s): <u>Hom</u> pression Syster tation(s): 3 posited: 2002-1 position Author zoff, E.D., Tain	e HSOD A4V mutar N19/pdb IDOREDUCTASE o sapiens n: Escherichia coli 0-16 Released: 2002-114 r(s): <u>Cardoso, R.M.F., The</u> er, J.A.	1 t -27 <mark>ayer, M.M., DiDonato, M</mark>	<u>., Lo, T.P., Bruns, C.K.</u> ,
3D View: Structure Ligand Interaction Standalone Viewers Protein Workshop Lig Global Symmetry: Cyc Global Stoichiometry:	and Explorer clic - C2 (3D : Homo 2-me	Nsity Exp Met Res R-V R-V View) r - A2	berimental Data thod: X-RAY DII solution: 1.86 Å Yalue Free: 0.26 Yalue Work: 0.20	a Snapshot FFRACTION 0 07	wPDB Validation	O 3D Report Full Report Percentile Ranks Value 0.209 10 0.3% 3.0% 3.9% setter at X-ray structures of similar resolution

SODC_DROME P61851 Superoxide dismutase [Cu-Zn] Created: today at 02:16 Summary Templates 50 Models Ŀ ΤX Template Results o Build Models 1 Quaternary Structure Sequence Similarity Alignment of Selected Templates Templates More -Clear Selection IfSort \$Name Title ♦GMQE ▼QSQE ♦Identity ♦Method ♦Oligo State ♦Ligands Coverage 3gtv.1.A Superoxide X-ray, 2.2Å homo-dimer 2 x ZN 2 🗸 0.80 0.94 60.26 dismutase [Cu-Zn] X-ray, 2.2Å homo-dimer 2 x ZN 2 🖤 \checkmark 3gtv.1.A Superoxide 0.82 0.93 61.33 dismutase [Cu-Zn] X-ray, 1.9Å homo-dimer 2 x ZN^C, ♥ 1n19.1.A Superoxide 0.81 0.92 61.18 Dismutase 2 x CU1 1 C [Cu-Zn] 4b3e.1.A SUPEROXIDE 0.79 0.91 61.18 X-ray, 2.1Å homo-dimer 2 x ZN^{IC}, ♥ DISMUTASE 2 x CU C [CU-ZN] 4b3e.1.A SUPEROXIDE 0.81 0.91 61.84 X-ray, 2.1Å homo-dimer 2 x ZN^{IC}, ♥ DISMUTASE 2 x CU [CU-ZN]

Iemplate Results e	
Templates Quaternary Structure Sequence Similarity Alignment of Selected Templates	Build Models 1
More -	Clear Selection
Target MVVKAVCVINGD AKGTVFFEQESSGTPVKVSGEVCGLAKGLHGFHVHEFGDNTNGGMSSGPHF 63 In19.1.A MATKVVVIGDD PVGI DFBCKSSTPVKVSGEVCGLAKGLHGFHVHEFGDNTAGGMSSGPHF 63 Target NPYGKEHGAPVDENRHLGDLGNIEATGDCPTKVNITDSKITLFGADSIIGRTVVHADADDLGCG 128 In19.1.A NPISSKHGAPVDENRHLGDLGNIEATGDCPTKVNITDSKITLFGADSIIGRTVVHADADDLGCG 128 In19.1.A NPISKHGGPKDERHUGDLGNIEATGDCPTKVNITDSKITLFGADSIIGRTVVHADADDLGCG 130 Target GHELSKSTGNAGARIGCGVIGIAKV 153 In19.1.A GYELSKTGNAG RIGCGVIGIAKV 153 In19.1.A GYELSKTGNAG RIGCGVIGIAKV 154	PV Cartoon C C

emplates Quaternary Structure Sequence Similarity Alignment of Selected Templates		Build Models (1)
lore -		Clear Selection
Target MVVKAVCVINGD AKGTVFFEQESSGTPVKVSGEVCGLAKGLHGFHVHEFGDNTNGCMSSGPHF n19.1.A MATKVVAVIGGGPVGIINFBQKESNGPVKV×GSIGLTEGLHGFHVHEFGDNTAGCTSAGPHF	63 65	
Target NPYGKEHGAPVDENRHLGDLGNIEATGDCPTKVNITDSKITLFGADSIIGRTVVVHADADDLGQG n19.1.ANPDSRKHGGPKDEERHVGDLGNVTÅDKDGVADVSIEDSVIČLSGDHÖJIGRTUVVHADADDLGKG	128	
Target GHELSKSTGNAGARIGCGVIGIAKV	153	
n19.1.A GNEESTKTGNAGSRUACGVIGDAQ-	154	
		That y
		₽V ▲ Cartoon ▲ 🗈 ► ▲











Cross validate with several other templates if possible



Cross validate with several other templates if possible

Template Results e	
Templates Quaternary Structure Sequence Similarity Alignment of Selected Templates	Build Models 5
More -	Clear Selection
Target MVVKAVCVINGDAKGTVFFEQESSGTPVKVSGEVCGLAKGLHGFHVHEFGDNTNGCMSSGP 3gtv.1.A -ATKAVCVIGGOGPVQGIINFEXESNGPVKVWGSTGLTEGLHGFHVHEFGDNTNGCTSAGA 3gtv.1.A -ATKAVCVIGG-DGPVQGINFEXESNGPVKVWGSTGLTEGLHGFHVHEFGDNTAGCTSAGA 3gtv.1.A - TKAVCVIGG-DGPVQGINFEXESNGPVKVWGSTGLTEGLHGFHVHEFGDNTAGCTSAGA 1n19.1.A MATKVVVIGG-DGPVQGINFEXESNGPVKVWGSTGLTEGLHGFHVHEFGDNTAGCTSAGA 319e.1.A MATKVVVIGG-DGPVQGINFEXESNGPVKVWGSTGLTEGLHGFHVHEFGDNTAGCTSAGA Target HFNPYGKEHGAPVDENRHLGDLGNIEATGDC-PTKVNITDSKITLFGADSIIGRTVVVHADADDL 3gtv.1.A FNPLSRKHGGPKDEERHVGDLGNIEATGDC-PTKVNITDSKITLFGADSIIGRTVVVHADADDL 3gtv.1.A FNPLSRKHGGPKDEERHVGDLGNIEATGDC-VANVSIEDKVI SGEHJIIGRTVVVADADDLL 3gtv.1.A FNPLSRKHGGPKDEERHVGDLGNIEATGDC-VANVSIEDKVI SGEHJIIGRTVVVADADDLL 3gtv.1.A FNPLSRKHGGPKDEERHVGDLGNIEATGDC-VANVSIEDKVI SGEHJIIGRTVVVADADDLL 3gtv.1.A FNPLSRKHGGPKDEERHVGDLGNIEATGDC-VANVSIEDKVI SGEHJIIGRTVVVADDDLL 3gtv.1.A FNPLSRKHGGPKDEERHVGDLGNIEATGDC-VANVSIEDKVI SGEHJIIGRTVVVADDDLL 3gtv.1.A FNPLSRKHGGPKDEERHVGDLGNVTOKDG-VANVSIEDKVI SGEHJIIGRTVVVADDDLL 3gtv.1.A FNPLSRKHGGPKDEERHVGDLGNVTOKDG-VANVSIEDKVI SGEHJIIGRTVVVADDDLL 3gtv.1.A FNPLSRKHGGPKDEERHVGDLGNVTOKDG-VANVSIEDKVI SGEHJIIGRTVVVADDDLL 3gtv.1.A FNPLSRKHGGPKDEERHVGDLGNVTOKDG-VANVSIEDKVI SGEHJIIGRTVVVADDDLL 3gtv.1.A FNPLSRKHGGPKDEERHVGDLGNVTOKDG-VANVSIEDKVI SLSGEHJIIGRTVVVADDDLL 3gtv.1.A FNPLSRKHGGPKDEERHVGDLGNVTOKDG-VANVSIEDKVI SLSGEHJIIGRTVVAADDDLL 3gtv.1.A FNPLSRKHGGPKDEERHVGDLGNVTOKDG-VANVSIEDKVI SLSGEHJIIGRTVVAADDDLL 3gtv.1.A FNPLSRKHGGPKDEERHVGDLGNVTOKDG-VANVSIEDKVI SLSGEHJIIGRTVVAADDDLL 3gtv.1.A FNPLSRKHGGPKDEERHVGDLGNVTOKDG-VANVSIEDKVI SLSGEHJIIGRTVVAADDDLL 3gtv.1.A FNPLSRKHGGPKDEERHVGDLGNVTOKDG-VANVSIEDKVI SLSGEHJIIGRTVVAADDDLL 3gtv.1.A FNPLSRKHGGPKDEERHVGDLGNVTOKDG-VANVSIEDKVI SLSGEHJIIGRTVVAADDDLL 3gtv.1.A FNPLSRKHGGPKDEERHVGDLGNVTOKDG-VANVSIEDKVI SLSGEHJIIGRTVVAADDDL	61 62 61 62 63 61 125 126 126 126 127 126
Target GQGGHELSKSTGNAGARIGCGVIGIAKV 3gtv.1.A GKGGNEESTKTGNAGSRLACGVIGIAKI 3ly.1.A GLGGNELSKTGNAGSRLACGVIGIAKI 3gtv.1.A GKGGNEESTKTGNAGSRLACGVIGIAKI	153 153 154 153
319e.1.A GLGGHELSKTTGNAGGRIACGVIG AKI	154 THR 95 A

Adjusting alignments with DeepView

Template Results e

Templates	Quaternary Structure	Sequence Similarity	Alignment of Selected Templates		Build Models 5
More 🗸					Clear Selection
Target MV	VKAVCVINGDAKGT	VFFEQESSGTPVKVSGEV	CGLAKGLHGFHVHEFGDNTNGCMSSO	P 61	
3gtv.1.A - A	TKAVCVL GDGPVQGI	INFEOKESNGPVKVWGSI	GLTEGLHGFHV DEFGDNTAGCTSAC	P 62	
319y.1.A MP	AKAVCVLRGDVSGT	VFFDODEKSPVVVSGEV	GLTKGKHGFHVHEFGDNTNGCTSAC	A 61	
3gtv.1.A	TKAVCVLKGDGPVQGI	INFEOKESNGPVKVWGSI	GLTEGLHGFHVHEFGDNTAGCTSAC	P 62	T
1n19.1.A MA	TKVVAVLCGDGPVQGI	INFEQKESNGPVKVWGSI	GLTEGLHGFHVHEFGDNTAGCTSAC	P 63	
319e.1.A M P	AKAVCVLRGDVSGT	VFFDOODEKSPVVVSGEV	GLTK <u>GKHGFHV</u> ÅEFGDNTNGC <u>TSA</u> C	A 61	
Target HF	NPYGKEHGAPVDENRHLG	DLGNIEATGDC-PTKVNI	TDSKITLFGADSIIGRTVVVHADADI	L 125	
3gtv.1.A DF	NPLSRKHGGPKDEERHVG	DLGNVTAGKDG-VANVSI	EDRVI SCEHSIIGRTMVV DEKODI	L 126	A ALLER N
319y.1.A	NPEKQDHGGPSSAVRHVG	DLGNIEAIEDAGVTKVSI	ODS OI DLHGPN SIIGRTLVV HADPDI	L 126	
3gtv.1.ABF	NPLSRKHGGPKDEERHVG	DLGNVTAGKDG-VANVSI	EDRVI SCEHSIIGRTMVV BEKODI	L 126	
1n19.1.A	NPLSRKHGGPKDEERHVG	DLGNVTADKDG-VADVSI	EDS VIDLSGDHDIIGRTLVVDE KADI	L 127	A DITON SA
319e.1.A	N P E K Q D H G G P S S A V R H V G	DLGNIEAIEDSGVTAVSI	ODSQISLHGPNSIIGRTLVVRADPDI	L 126	
Target GO	GGHELSKSTGNAGARIGC	GVIGIAKV C	econdary structure	153	The second se
3gtv.1.AGK	GGNEESTKTGNAGSRIAC	aviana-	econdary structure	153	A A A A A A A A A A A A A A A A A A A
319v.1.A GL	GGNELSKTTGNAGGRIAC	GVIG AKI İS	s in good agreement!	154	
3gtv.1.A GK	GGNEESTKTGNAGSRLAC	GUIGDA -		153	
1n19.1.A GK	GGNEESTKTGNAGSRLAC	GUIGDAO-		154	
319e.1.A GL	GGHELSKTTGNAGGRIAC	GVIGLAKI		154	THR 95 A
		V			
					🚆 PV 🔺 Cartoon 🔺 🔯 🕨 📥 🚭
					3gtv.1.A 🛛 🗶

Adjusting alignments with DeepView



https://spdbv.vital-it.ch/

Adjusting alignments with DeepView



Viewing Ramachandran plot in deepview



Viewing Ramachandran plot in deepview



- Green crosses indicates residues with *favored dihedral angles*
- Red cross indicated residues with *unfavored* dihedral angles
- If most of these are in loop regions that is acceptable

Validating your model - SAVES

The Structure Analysis and Verification Server version 4

[SAVES] XdVal MTZdump Ramachandran Plot pdbU pdbSNAFU (Check for ADIT compliance) PROCHECK Verify3D ERRAT

This metaserver runs 6 programs for checking and validating protein structures during and after model refinement.

This server processed 3,677.3 jobs per month in the last 12 months. See more usage statistics here

Monthly totals plot

- Checks the stereochemical quality of a protein structure by analyzing residue-by-residue geometry and overall structure PROCHECK geometry. [Reference]
- WHAT_CHECK Derived from a subset of protein verification tools from the WHATIF program (Vriend, 1990), this does extensive checking of many sterochemical parameters of the residues in the model. [Reference]
 - Analyzes the statistics of non-bonded interactions between different atom types and plots the value of the error function versus ERRAT position of a 9-residue sliding window, calculated by a comparison with statistics from highly refined structures. [Reference]
- Determines the compatibility of an atomic model (3D) with its own amino acid sequence (1D) by assigning a structural class VERIFY 3D based on its location and environment (alpha, beta, loop, polar, nonpolar etc) and comparing the results to good structures. References: [Bowie et al., 1991; Luethy et al., 1992].
- Calculates the volumes of atoms in macromolecules using an algorithm which treats the atoms like hard spheres and calculates a statistical Z-score deviation for the model from highly resolved (2.0 Å or better) and refined (R-factor of 0.2 or better) PDB-PROVE deposited structures. [PUBMED Reference].
- CRYST1 record We take the CRYST1 record and search the entire PDB for matches and report these as possibly similar structures.

Ramachandran We produce an interactive Ramachandran plot. Also a standalone server linked above. Plot

http://services.mbi.ucla.edu/SAVES/

Summary

- Final model is only as good as the input template.
- Identity is usually the best benchmark for final quality but don't ignore SS prediction.
- Always remember to cross check, more models means more confidence (for the most part).
- Try different model tools (my preference are I-TASSER and robetta but they are slow)
- You may need to try several different alignments when homology is low – Alignment is the place most likely to introduce errors.
- Always validate! Does your model make chemical sense?

Further reading

- For more on definitions of protein structure:
 - Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963). JMol Biol 7:95-99.
 - Richardson. (1981). The Anatomy and Taxonomy of Protein Structure, Advances in protein chemistry; 34. 167 -339 (New York: Academic Press).
 - Kuhlman, B., and Baker, D. (2000). Native protein sequences are close to optimal for their structures. Proceedings of the National Academy of Sciences 97, 10383–10388.
 - Zhang, Y. (2009). Protein structure prediction: when is it useful? Current Opinion in Structural Biology 19, 145–155.

Further reading

For more on folding and algorithms:

- Li, Z., and Scheraga, H.A. (1987). Monte Carlo-minimization approach to the multiple-minima problem in protein folding. Proceedings of the National Academy of Sciences 84, 6611–6615.
- Simons, K.T., Kooperberg, C., Huang, E., and Baker, D. (1997). Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and bayesian scoring functions1. Journal of Molecular Biology 268, 209–225.
- Jones, D.T. (1999). Protein secondary structure prediction based on positionspecific scoring matrices1. Journal of Molecular Biology 292, 195–202.
- Song, Y., DiMaio, F., Wang, R.Y.-R., Kim, D., Miles, C., Brunette, T., Thompson, J., and Baker, D. (2013). High-Resolution Comparative Modeling with RosettaCM. Structure 21, 1735–1742.

Further reading

For more on Swiss model in particular:

- Benkert, P., Biasini, M., and Schwede, T. (2011). Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 27, 343–350.
- Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T.G., Bertoni, M., Bordoli, L., et al. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Research 42, W252–W258.
- Bienert, S., Waterhouse, A., de Beer, T.A.P., Tauriello, G., Studer, G., Bordoli, L., and Schwede, T. (2017). The SWISS-MODEL Repository—new features and functionality. Nucleic Acids Research 45, D313–D319.