MICROBIOLOGY INDONESIA
ISSN 1978-3477

Volume 3, Number1, April 2009
p 47-50

T-cell Epitopes of *Mycobacterium tuberculosis* Antigen 85 Complex Potential for Generating Antibody: an Immunoinformatics Study

YOGY SIMANJUNTAK

Microbiology Division, Research Center for Biology, Lembaga Ilmu Pengetahuan Indonesia,
Jalan Raya Cibinong KM 46, P.O Box 25, Cibinong 16911, Indonesia
Phone: +62-21-8765066, Fax: +62-21-8765062, Email: yogy_simanjuntak@yahoo.com

The present study was conducted to predict T-cell epitopes of the antigen 85 complex capable of stimulating antibody generation by using immunoinformatics approaches. By applying computational biology software, the available data of the antigen 85 complex and related-epitopes would be turned into more constructive and useful scientific informations for the development of multiepitope-based anti-TB vaccine. The identification of T-cell epitopes capable of generating antibody was done by the 3DEX program, discotope analysis and PyMol program. Selected peptides having individual amino acids localized on the predicted antibody-binding sites were subjected to antigenic property analysis, including their hydrophilicity, flexibility and antigenic propensity. The 3DEX program identified 17 peptides having at least four individual amino acids located on the antigen surface. However, after homology analysis with preselected distance of 7 Å and taking into account the spatial neighborship, only seven peptides of antigen 85A, 85B and 85C (3, 3 and 1 peptide(s) respectively) had individual amino acids overlapping the predicted antibody-binding site. Peptides 17838, 21780, 21275 and 36131 had an average score of antigenic propensity above 1.0. In conclusion, there are seven peptides representing T-cell epitope of antigen 85 complex that could potentially be capable of generating an antibody response. The seven peptides, P17838, P21093, P36131, P21275, P21780, P21796 and P10839, are suitable candidates for further study in order to develop a subunit-based multiepitope anti-TB vaccine.

Key words: *Mycobacterium tuberculosis*, multiepitope anti-TB vaccine, T-cell epitope, antibody, immunoinformatics

Tuberculosis (TB) is an infectious disease caused predominantly by the pathogenic bacteria called, *Mycobacterium tuberculosis*. Tuberculosis is still a major health problem in Indonesia and considered as the third main cause of death after cardiovascular and respiratory diseases. As the third country in descending order of TB case numbers after India and China, Indonesia has to fight the disease seriously (http://www.who.int/gtb.html). Nowadays, wide use of rifampicin containing regimens often leads to a steady increase in multidrug resistant tuberculosis (Kim 2004). It is appealing to speculate that antibiotic therapy becomes a two-edged sword in TB intervention. Furthermore, failure of the BCG vaccine to protect in endemic regions calls for the development of effective vaccines.

Instead of using whole protein preparations as vaccine candidates, researchers currently consider that T-cell epitopes may be used in the development of vaccines. Some investigators are trying to identify the kind of peptide antigens that could evoke not only T-cell mediated response, but also antibody synthesis (Sarhan *et al*. 2007). Such a study is quite tricky, taking into consideration antibody designated to recognize surface epitope of the antigen, while in contrast T-cell epitopes are more hydrophobic (Laver *et al*. 1990).

The antigen 85 complex, as the most common proteins in *M. tuberculosis* culture fluids, would be suitable for the purpose of interest. This antigen is a strongly immunogenic, stimulating not only T-cell-mediated reaction but also humoral immune response. The representing constituents, 85A, 85B and 85C, are encoded by three genes located at different sites in the mycobacterial genome. All are a fibronectin-binding protein and strongly immunogenic both in natural and experimental studies stimulating antibody synthesis and T-cell-mediated reactions (Wiker and Harboe 1992).

Fortunately, more than a hundred T-cell epitopes of the antigen 85 complex has been reported in the Immune Epitope Database and Analysis Resource (www.immuneepitope.org). However, their potency to induce antibody generation remains unclear. In view of therapeutic vaccine development that could evoke both cell-mediated and humoral immune response, an immunoinformatics study has to be done prior to experimental studies. This avoids labor-intensive, costly and time-consuming works. Therefore, this study was conducted to predict T-cell epitopes of the antigen 85 complex capable of stimulating antibody generation using an immunoinformatics approaches. By applying computational biological software, the available data of the antigen 85 complex and related-epitopes can be turned into more constructive and useful scientific information for the development of multiepitope-based anti-TB vaccine.

MATERIALS AND METHODS

Data Sources. Peptide sequences representing T-cell epitopes of the antigen 85 complex were derived from many reports in the Immune Epitope Database and Analysis Resource (www.immuneepitope.org), determined by various assays including ELISA, cytokine bioassay, ELISPOT, 51 Chromium release and proliferation methods. The pdb files of the antigen 85 complex corresponding to known structures of antigen 85A (1SFR), 85B (1F0N) and 85C (1DQZ) were retrieved from the RCSB Protein Data Bank (www.rcsb.org).

Screening of T-Cell Epitopes. Three databases consisting of 50, 57 and 25 different peptide sequences representing T-cell epitopes were analyzed for their surface structure against respective antigen 85A, 85B and 85C. The surface scan implemented in the 3D-Epitope-Explorer (3-DEX) program was performed to screen the surface structure of each peptide. The probability of amino acids to be on surface exposure was set to equal or greater than 50% whilst joker function was activated (Schreiber *et al*. 2005). Peptide

sequences were selected for further analysis if more than three individual amino acids were likely on the surface exposure of the respective antigen.

Identification of T-Cell Epitopes Capable of Generating Antibody. By using the 3D-Epitope-Explorer program, the selected peptides were subjected to homology analysis against their respective structural pdb files. Anticipating a conformational epitope, a distance of 7 Å (Angström) was preselected with the highest possible frame size. The previous result of surface scan was used to determine the most suitable homologous hit. Antibody-binding sites of respective antigens were predicted by using discotope analysis (Andersen *et al*. 2006). The peptide sequence and antibody binding sites could then be highlighted in the pdb structure file using the program PyMol.

Analysis of Antigenic Properties. Selected peptides with individual amino acids overlapping the predicted antibody-binding sites were subjected to antigenic property analysis including hydrophilicity, flexibility and antigenic propensity as described by Karplus and Schulz (1985); Parker *et al*. (1986) and Kolaskar and Tongaonkar (1990).

RESULTS

Three databases consisting of 50, 57 and 25 different T-cell epitopes were analyzed for their surface structure against respective antigen 85A, 85B and 85C. The 3DEX software identified that 5 out of 55 peptides have more than three individual amino acids located on surface area of 3-dimensional (3D) structure-antigen 85A. As many as 11 peptides possessed more than three individual amino acids that were exposed on the surface of antigen 85B. Only a single peptide representing T-cell epitope of antigen 85C was considered for further analysis (Table 1).

Table 1 Screening of surface structure of T-cell epitope against respective antigen*

Variable	Number of amino acid accessible on surface area		
	None	1-3	4 ≤
Antigen 85 A (n=50)	12	33	5 (FSRPGLPVEYLQVPSPSMGR; GLRAQDDFSGWDINTPAFEW; GPKEDPAWQRNDPLLNVGKL; LGGNNLPAKFLEGFVRTSNI; WDINTPAFEWYDQSGLSVVM)
Antigen 85 B (n=57)	16	30	11 (GMGPSLIGL; GPSLIGLAM; SMAGSSAMIL; FLTSELPQWL; IGLSMAGSSAMILAAY; FLTSELPQWLSANRAVKP; GGYKAADMWGPSSDPAWE; GPSSDPAWERNDPTQQIP; RNDPTQQIPKLVANNTRL; WYSPACGKAGCQTYKWET; LQVPSPSMGRDIKVQFQSGG)
Antigen 85 C (n=25)	8	16	1 (DWYQPSQSNGQNYTYKWETF)

* values are expressed as the total number

Table 2 Homology and surface scan analysis of the selected peptide sequence against the 3D- structure of antigen 85A [1SFR]*

Peptide ID	Distance [Angströms]: 7 Scan surface: Threshold ≥ 50% Max. number of jokers: 2 Reference atom: Alpha-C atom
	[1SFR_85A]
P17838	FSRPGLPVEYLQVPSPSMGR **PHE 1|SER 2|ARG 3|PRO 4|GLY 5|LEU 6**|PRO 7|VAL 8|GLU 9| TYR 10|LEU 11 |GLN 12|VAL 13|**PRO 14**|SER 15|**PRO 16**|SER 17|MET 18|GLY 19| ARG 20
P21093	GLRAQDDFSGWDINTPAFEW LEU 42|**ARG 43**|ALA 44|GLN 45|**ASP 46|ASP 47|PHE 48**|SER 49|GLY 50 | TRP 51|ASP 52|ILE 53|ASN 54|THR 55|PRO 56|ALA 57|PHE 58|GLU 59|TRP60
P36131	LGGNNLPAKFLEGFVRTSNI LEU 219|GLY 220|GLY 221|**ASN 222|ASN 223**|LEU 2 24|PRO 225|ALA 226| LYS 227|PHE 228|LEU 229|GLU 230
P21670	GPKEDPAWQRNDPLLNVGKL GLY 181|PRO 182|**LYS 183**|GLU 184|ASP 185|PRO 186|ALA 187|TRP 180
P72312	WDINTPAFEWYDQSGLSVVM TRP 51|ASP 52|ILE 53|ASN 54|THR 55|PRO 56|ALA 57|PHE 58| **GLU 59**| TRP 60|TYR 61|ASP 62|GLN 63|SER 64|GLY 29|LEU 66|SER 67|VAL 68|VAL 69| MET 70

*underlines are homolog amino acid sequences; amino acids which are surface exposed are marked in bold

Table 2 shows the homology and surface scan profile of five selected peptides against the 3D-structure antigen 85A (chain A). The 3DEX software identified two peptides, P17838 and P72312, which were completely aligned within the preselected distance of 7 Å against antigen 85A. However, the last contiguous peptide (P72312) in addition to peptide 21670 had only one individual amino acid located on the surface of antigen 85A (E59 and K183, respectively). Although not perfectly homologous within the preselected distance, peptide P17838 and P21093 had at least 3 contiguous individual amino acids exposed on the surface of antigen 85A (F1 S2 R3 P4 G5 L6 and D46 D47 F48, respectively). Interestingly, in addition to discontinuous amino acids P14 and P16, three contiguous amino acids (R3 P4 G5) of peptide P17838 overlapped the predicted antibody-binding site at the neutralizing face of antigen 85A (Fig 1). Four individual amino acids within the peptide P21093 were located on the surface of antigen 85A (R43 D46 D47 F48). However, only two continuous amino acids (D46 D47) overlaid the predicted antibody-binding site. In addition, two contiguous amino acids (N222 N223) within the peptide P36131 were located in the region of the predicted antibody-binding site.

The 3DEX software identified 5 out of 11 selected peptides (namely: SMAGSSAMIL, FLTSELPQWL, IGLSMAGSSAMILAAY, FLTSELPQWLSANRAVKP and RNDPTQQIPKL VANNTRL) that were perfect homolog within the preselected distance of 7 Å against the 3D-stucture of the antigen 85B. However, most of the homologous hits did not fit with the surface scan analysis. Only six selected peptides having surface structure are described in Table 3. Although partly aligned within the preselected distance against antigen 85B, three peptides (P21275, P21780 and P21796) had at least two individual amino acids overlapping the predicted-antibody binding site (Fig 2). Some amino acids within three peptides (P73294, P54977 and P20101) were also found to be part of antigen surface area. However, all of these were independent of the predicted antibody-binding site.

Table 4 shows homology and surface scan analysis of peptide P10839 against the 3D- structure of antigen 85C (chain B). The peptide was partly homologous within the preselected distance of 7 Å against antigen 85C. The amino acids, Q582 Q585 S586 G588 N590, were localized on the surface exposure of the antigen. Interestingly, all of those amino acids overlapped the predicted antibody-binding site (Fig 3).

Antigenic properties of seven peptides are presented in Table 5. The seven peptides had at least two individual amino acids overlapping the predicted antibody-binding site. Four peptides (P17838, P21780, P21275 and P36 131) had an average score of antigenic propensity above 1.0. Three peptides (P10839, P21796 and P21093) were considered the most hydrophilic compared to the others.

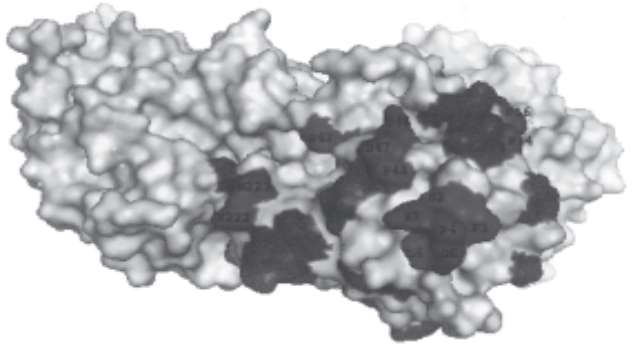

Fig 1 Individual amino acids located on the surface antigen 85A 3D structure within the three peptides (**F1S2R3P4G5L6**PVEYLQV **P14SP16**SMGR, GL**R43**AQ**D46D47F48**SGWDINTPAFEW, LGG**N2 22N223**LPAKFLEGFVRTSNI). The predicted antibody-binding site of antigen 85A (Chain-A) is shown in blue. Amino acids of the respective peptides overlap the predicted-antibody binding site are shown in magenta. Individual amino acids independent of the predicted-antibody binding site are shown in red.

DISCUSSION

The only widely used vaccine against childhood TB, the BCG vaccine, is unlikely to have a significant impact on an adult TB epidemic. Therefore, the domain of the anti-TB vaccine development has been explored extensively in order to find out the most efficient prophylactic vaccines. The new techniques for preparing anti-TB vaccine include development of DNA vaccine, modified BCGs and multiepitope-based vaccines (Sarhan 2007). The last mentioned technique becomes relatively straightforward with the availability of proteomic tools such as peptide synthesizer machine.

New vaccines being developed against TB focuses on the T lymphocyte since it is the central protection against TB (Sarhan 2007). Interestingly, a study reported by Spouge *et al*. (1987) revealed that strong conformational propensities enhance T-cell antigenicity. Conformational epitopes are required for neutralizing antibody, thus they have to be located on the surface of a given antigen. In Immune Epitope Database and Analysis Resource, more than a hundred peptides representing T-cell epitopes of the antigen 85 complex has been reported by various studies. The question now is which peptide(s) will be potentially capable to stimulate antibody response?

Table 3 Homology and surface scan analysis of the selected peptide sequence against the 3D- structure of antigen 85B [1F0N]*

Peptide ID	Distance [Angströms]: 7 Scan surface: Threshold ≥ 50% Max. number of jokers: 2 Reference atom: Alpha-C atom
	[1F0N_85B]
P21275	**GM**GPS**L**IGL **GLY 158|MET 159|**GLY 160|PRO 155|SER 156
P21780	**GPSL**IGLAM **GLY 5|PRO 4|SER 2|**LEU 6
P21796	G**PSSDP**AWERNDPTQQIP GLY 181|**PRO 182|SER 183|SER 184|**ASP 185|**PRO 186|**ALA 187|TRP 180
P73294	W**YSP**ACG**KA**G**CQ**TYKWET TRP 180|**TYR 83|**SER 84|**PRO 85|**ALA 86|CYS 87|GLY 88|**LYS 89|ALA 90|** GLY 91|**CYS 92|**GLN 93
P54977	RNDPTQQI**PK**LVANNTRL ARG 190|ASN 191|ASP 192|PRO 193|THR 194|GLN 195|GLN 196|ILE 197|**PRO 198|LYS 199|**LEU 200|VAL 201|ALA 202|ASN 203|ASN 204|THR 205|ARG 206|LEU 207
P20101	GGY**K**A**A**DMWGPSSDPAWE GLY 88|GLY 172| TYR 174|**LYS 175|**ALA 176|**ALA 177|**ASP 178|MET 179|TRP 180|GLY 181|PRO 182|SER 183|SER 184|ASP 185|PRO 182|ALA 177|TRP 180

*underlined are homolog amino acid sequences; amino acids which are surface exposed are marked in bold.

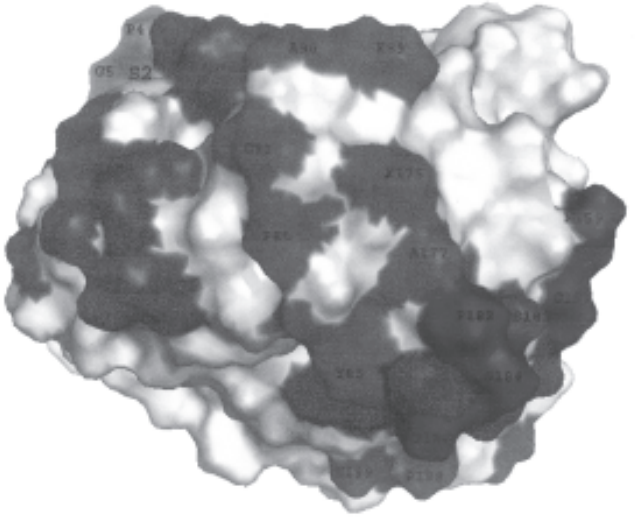

Fig 2 Individual amino acids located on the surface antigen 85B 3D-structure within the six peptides (**G158M159**GPSLIGL, **G5P4S2** LIGLAM, G**P182S183S184**D**P186**AWERNDPTQQIP, W**Y83**S **P85**ACG**K89A90**G**C92**QTYKWET, RNDPTQQI**P198K199**LV ANNTRL and GGY**K175** A**A177**DMWGPSSDPAWE). The predicted antibody-binding site of antigen 85B is shown in blue. Amino acids of the respective peptides which overlap the predicted antibody-binding site are shown in magenta. Individual amino acids independent of the predicted antibody-binding site are shown in red.

Table 4 Homology and surface scan analysis of the selected peptide sequence against the 3D structure of antigen 85 [1DQZ]*

Peptide ID	Distance [Angströms]: 7 Scan surface: Threshold ≥ 50% Max. number of jokers: 2 Reference atom: Alpha-C atom
	[1DQZ_85C]
P10839	D**W**Y**QP**S**QS**N**G**Q**N**YTYKWETF ASP 579 $_{683}$|TRP 580 $_{678}$|TYR 581|**GLN 582|**PRO 583|SER 584|**GLN 585|SER 586|**ASN 587|**GLY 588|**GLN 589|**ASN 590**

*underlined are homolog amino acid sequences; amino acids which are surface exposed are marked in bold; subscript indicates an alternative amino acid.

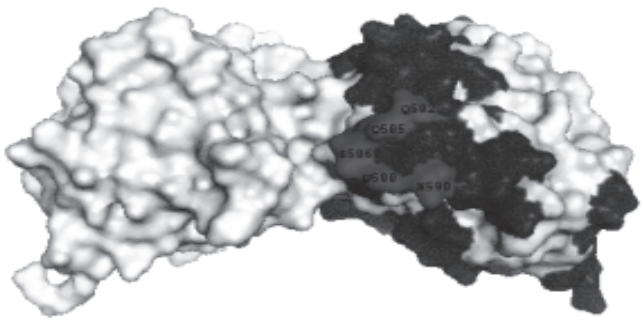

Fig 3 Individual amino acids located on the surface antigen 85C 3D-structure within the selected peptide (DWY**Q582**PS**Q585S586** N**G588**Q**N590**YTYKWETF). The predicted antibody-binding site of antigen 85C (Chain-B) is shown in blue. Amino acids of the respective peptides which overlap the predicted antibody-binding site are shown in magenta.

Table 5 Antigenic properties of selected peptides*

Peptides	Antigen	Hydrophilicity	Flexibility	Antigenic propensity
P17838	85A	0.746 (-1.200 to +3.271)	0.994 (0.941 to 1.062	1.088 (0.961 to 1.158)
P21093	85A	2.147 (0.286 to 4.443)	1.007 (0.979 to 1.030)	0.960 (0.924 to 1.004)
P36131	85A	0.437 (-1.943 to 2.914)	1.002 (0.982 to 1.052)	1.017 (0.954 to 1.071)
P21275	85B	-0.438 (-0.914 to -0.200)	1.010 (0.999 to 1.021)	1.028 (1.007 to 1.068)
P21780	85B	-1.557 (-2.329 to -0.914)	0.966 (0.949 to 0.983)	1.075 (1.061 to 1.095)
P21796	85B	3.957 (2.757 to 6.043)	1.049 (1.001 to 1.103)	0.946 (0.890 to 1.012)
P10839	85C	3.571 (0.400 to 6.386)	1.064 (0.977 to 1.148)	0.969 (0.926 to 1.025)

*values are expressed as average score (minimum to maximum).

In present study, the validated-3DEX program only identified a few number of T-cell epitopes that potentially mimic such conformational epitopes which have individual amino acids located on surface of the antigen 85 complex. It seems the characteristic discrepancy between T-cell and B-cell epitopes is the main factor. The antibody is designated to recognize an epitope located on the surface of an antigen. Generally, this kind of epitope consists of hydrophilic amino acids. In contrast, T-cell epitopes are mostly hydrophobic. It was found in the present study that only seven peptides (P17838, P21093, P36131, P21275, P21780, P21796 and P10839) have individual amino acids overlapping the predicted antibody-binding site of the respective antigen 85. This means that in addition to a cell-mediated immune response, these seven peptides could potentially stimulate antibody generation either directly or indirectly, through major histocompatibility complex (MHC) class II.

The newly-published peptide derived from *M. tuberculosis* Rv1490 surface protein was proposed as the potential vaccine candidate taking the same item for multiepitope-based vaccine development (Patarroyo *et al*. 2008). This comprehensive research involved a bioinformatics approach, in an *in vitro* and *in vivo* study, suggesting two peptides AEILVKYAQLADKRARVYVL (11 060) and FGRVESHADYHDWVCEHVTP (11 073) play an important role in TB pathophysiology. The antigenic propensity of peptide 11 060, based on the finding of Kolaskar *et al.* (1990) is 1.072, while peptide 11 073 is 1.067. Interestingly, these average scores are considered lower compared with the P17838 and P21780 scores (1.088 and 1.075 respectively). This antigenic propensity serves as a gold standard in predicting antigenic determinant. Antigenic propensity combines hydrophilicity, flexibility and surface accessibility scores all at the same time.

Ultimately, it would be very challenging to conduct a confirmation study, such as immunization of experimental animals to reveal whether these seven peptides are indeed capable of strongly evoking both the T-cell mediated immune and the humoral-immune responses.

In conclusion, there are seven peptides representing T-cell epitope of the antigen 85 complex that could potentially be capable of generating an antibody response. The seven peptides, P17838, P21093, P36131, P21275, P21780, P21796 and P10839, are strong candidates proposed for further study in order to develop a subunit-based multiepitopic anti-TB vaccine.

ACKNOWLEDGEMENT

The author thanks to Andreas Schreiber (Institute for Biomedical Research-Frankfurt/Main, Germany) for providing a bioinformatics program, 3D-Epitope-Explorer. The author also thanks Okta Rena for editorial work.

REFERENCES

Andersen PH, Nielsen M, Lund O. 2006. Prediction of residues in discontinuous B cell epitopes using protein 3D structures. Prot Sci 15:2558-67.

Karplus PA, Schulz GE. 1985. Prediction of chain flexibility in proteins-a tool for the selection of peptide antigens. Naturwissenschaften 72:212-3.

Kim SJ. 2004. Second-line drug susceptibility testing: where are we and where are we going? Int J Tuberc Lung Dis 8:S6-7.

Kolaskar AS, Tongaonkar PC. 1990. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett 276:172-4.

Laver WG, Air GM, Webster RG, Smith-Gill SJ. 1990. Epitopes on protein antigens: Misconcptions and realities. Cell 61:553.

Parker JM, Guo D, Hodges RS. 1986. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. Biochem 25:5425-32.

Patarroyo MA, Curtidor H, Plaza DF, Ocampo M, Reyes C, Saboya O, Barrera G, Patarroyo ME. 2008. Peptides derived from *Mycobacterium tuberculosis* Rv1490 surface protein implicated in inhibition of epithelial cell entry: Potential vaccine candidates? Vaccine 26:4387-95.

Sarhan MAA. 2007. Progress in tuberculosis vaccine development. Res J Med Med Sci 2:35-41.

Schreiber A, Humbert M, Benz A, Dietrich U. 2005. 3D-Epitope-Explorer (3DEX): Localization of conformational epitopes within three dimensional structures of proteins. J Comput Chem 26:879-87.

Spouge JL, Guy HR, Cornette JL, Margalit H, Cease K, Berzofsky JA, DeLisi C. 1987. Strong conformational propensities enhance T-cell antigenicity. J immunol 138:204-12.

Wiker HG, Harboe M. 1992. The antigen 85 complex: a major secretion product of *Mycobacterium tuberculosis*. Microbiol Rev 56:648-61.